Molecular Investigations of Iduronate-2-sulfatase Mutants

LAU Kin Chong

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Philosophy
in
Chemical Pathology

August 2006

Abstract

Abstract of thesis entitled:

Molecular investigations of iduronate-2-sulfatase mutants

Submitted by Lau Kin-Chong

for the Degree of Master of Philosophy in Chemical Pathology

at The Chinese University of Hong Kong in June 2006

In Hong Kong, we have four rare subtypes of Mucopolysaccharidoses (MPS). Only Mucopolysaccharidosis type II (MPS II) is the subtype that is X-linked. Deficiency of iduronate-2-sulfatase (IDS) leads to intralysosomal accumulation of undegraded glycosaminoglycans (GAG) and causes the phenotypes. Clinically, a broad spectrum of phenotypes is observed among the MPS II patients. This heterogeneity is partially contributed by the different mutations of the *IDS* gene. Molecular investigations of a specific disease-causing mutation can lead to a better understanding of the pathological mechanism of this disease.

In this study, direct sequencing techniques were used to identify the mutations in four Hong Kong Chinese MPS II patients. Three different mutations were found in exon 8 of the *IDS* gene. Two of them are novel which include a nonsense mutation at

codon 369 (S369X) and a missense mutation at codon 339 (L339P). The known mutation was a nonsense mutation at codon 389 (Q389X). Three female carriers were detected and prenatal diagnosis was performed in one family. The fetus was finally confirmed to be normal.

To study possible genotype-phenotype correlation, in-situ mutagenesis studies at the codon 339 were carried out by cell-based *in vitro* expression. Transient expression of L339P and L339R in COS-7 cells revealed that both mutants had similar mRNA expression level as the wild-type. However, the enzymatic activities were found to be significantly reduced (less than 2.5 %). Additionally, structural analysis suggested that Leu339 residue is close to substrate binding site of the enzyme. In this study, wild-type IDS was also expressed using five different cell-free *in vitro* expression systems. Using cell-free *in vitro* expression systems, IDS precursors were successfully synthesized either based on prokaryotic or eukaryotic translation machinery. Results of cell-free *in vitro* expression and functional study of mutant IDS with modified N-terminal suggested that synthesis of catalytic active IDS may depend on C-terminal cleavages of glycosylated polypeptides.

摘要

目前在香港有四種罕有的黏多醣貯積症（簡稱黏多醣症）的病人，其中黏多醣症第二型是唯一一種性聯隱性基因遺傳病。 這些患者體內缺少一種酵素叫醛糖酸鹽硫酸脂脢，引致細胞內的黏多醣無法順利進行新陳代謝，而在細胞內異常地累積，損害各個不同的器官。 患者的臨床表徵差異甚大，是由於有多種不同的基因序列突變發生在醛糖酸鹽硫酸脂脢的基因上，導致對其酵素表達、酵素穩定性、酵素活性等，產生不同程度的影響。 進一步研究這些致病的基因突變，有助了解缺陷基因如何影響患者。

在本研究中，通過直接檢查患者的基因序列，成功在四位黏多醣症第二型的本地患者，確認了他們的醛糖酸鹽硫酸脂脢基因分別帶有三種不同的點突變。 這三種突變都是發生在經常出現點突變的第八個外顯子。 當中，在鹼基位置 serine369 及 leucine339 的核甘酸序列變異(S369X & L339P) ，是兩個新發現的突變基因未曾被文獻報導，只有在鹼基位置 glutamine389 發現的無義突變是已知的。 利用患者的缺陷基因，更成功檢測了三位女性帶因者，及有效地替其中一位帶因者進行產前檢查，確實胎兒並沒有遺傳該缺陷基因。

爲了進一步探討在鹼基位置 leucine339 發現的核甘酸序列變異對基因表達、酵素活性的影響，本研究利用了體外系統表達突變體，希望藉此聯結缺陷基因與臨床表徵的可存在關係。 通過 COS-7 猿猴細胞作瞬時表達，進行取代突變研究，兩種在鹼基位置 leucine339 的誤意突變 (L339P 和 L339R) 雖表達與野生型相近的 mRNA 量，但其酵素活性功能則大幅降低 (少於野生型的 2.5 %) 。 從本研究結果及三圍立體結構分析，估計鹼基位置 leucine339 可能接近一組特別的鹼基序列，負責將底物與醛糖酸鹽硫酸脂胺結合。 另外，本研究也試用了五個無細胞系統表達野生型的醛糖酸鹽硫酸脂胺。 結果發現，不論使用原核生物的轉譯技術或真核生物的轉譯技術，無細胞系統只能表達前驅蛋白。 此外，通過表達有氨基末端修飾的突變體和對它的功能分析，結果顯示具催化功能的醛糖酸鹽硫酸脂胺必須先經過醣化作用，及後完成羧基末端的蛋白質裂解，否則會完全失去酵素活性。

Acknowledgements

I would like to express my heartfelt thanks to the mighty God for granting me the strength and wisdom throughout my two-year M.Phil. study.

I am grateful to acknowledge my supervisor, Dr. Eric CW Lam (Department of Chemical Pathology, The Chinese University of Hong Kong) for setting realistic and firm goals of the project. And he gave me appropriate credit for having ideas. His critical thinking and innovative mind are the most that I sincerely admired.

I would like to give my special gratitude to Ms. Shirley SF Tong for her helpful technical guidance, especially on DNA sequencing; Ms. Priscilla TY Law for her valuable discussion and technical support in tissue culturing and protein expression. I would also like to thank all other colleagues in Department of Chemical Pathology.

Last but not least, I wish to express my thankfulness to you, the reader of this thesis. Thanks for sharing your time in my thesis.

Abby KC Lau

June 2006

Table of Contents

List of Tables

List of Figures

Appendices

Abbreviations

Short forms	**Full terms**
3D	3 dimension
4-Mu	4-methylumbelliferone
4-Mu-αIdu-2S	4-methylumbelliferyl-α-L-iduronide-2-sulphate.Na2
APS	ammonium persulphate
ATP	adenosine triphosphate
BGH	bovine growth hormone
BMT	bone marrow transplantation
bp	base pair
BSA	bovine serum albumin
C-terminal	carboxyl-terminal
cDNA	complementary deoxyribonucleic acid
CMV	cytomegalovirus
CV	chronic villi
dH_2O	deionized water
DMEM	Dulbccco's modified Eagle's medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
E. coli	*Escherichia coli*
ER	endoplasmic reticulum
ERT	enzyme replacement therapy
FBS	fetal bovine serum
FGE	formylglycine generating enzyme
FGly	formylglycine
g	gram
GAG	glycosaminoglycans
GFP	green fluorescent protein
GUS	glucuronidase
HBSS	Hank's balanced salt solution
IDS	iduronate-2-sulfatase protein
IDS	iduronate-2-sulfatase gene
IDUA	alpha-L-iduronidase protein

Short forms	Full terms
kb	kilobase pairs
kDa	kilodalton
L	litre
LAMP-1	lysosomal membrane-associated protein
LB	Lennox L
LEBT	lysosomal enzymes from bovine testis
LSDs	lysosomal storage diseases
m	milli
M	Molar
μ	micro
M6P	mannose-6-phosphate
MDE	mutation detection enhancement
min	minute(s)
mol	mole
MPS	Mucopolysaccharidoses
MPS I	Mucopolysaccharidosis type I
MPS II	Mucopolysaccharidosis type II
mRNA	messenger ribonucleic acid
MW	molecular weight
n	nano
N-terminal	amino-terminal
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylenglycol
RBS	ribosome binding site
RNA	ribonucleic acid
RNase H	ribonuclease H
rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse-transcriptase polymerase chain reaction
SDS	sodium dodecyl sulphate
sec	second(s)
SSCP	single-strand conformation polymorphism
SUMF 1	sulfatase modifying factor 1
TAE	Tris-acetate-EDTA

Short forms	Full terms
TEMED	N, N, N', N'-tetramethylenediamine
U	unit
WBC	white blood cells
WT	wild-type
Xa	active allele
Xi	inactive allele

Chapter 1 Introduction

1.1 Mucopolysaccharidosis type II as a lysosomal storage disease

Mucopolysaccharidoses (MPS) is a group of inborn errors of metabolism with defects in degradation of glycosaminoglycans (GAG) or mucopolysaccarides, in lysosomes. Up to date, there are 11 known enzyme deficiencies that give rise to 7 distinct MPS. Mucopolysaccharidosis type II (MPS II), MIM entry: #309900, is caused by the deficiency of the lysosomal enzyme iduronate-2-sulfatase (IDS) which initiates the stepwise degradation of GAG [Neufeld EF and Muenzer J, 2001]. The genetic defects lying in the *IDS* gene result in lysosomal accumulation of GAG. This disorder was first described by Charles Hunter in 1917 and has the eponym Hunter syndrome. In majority of lysosomal storage diseases (LSDs) including most MPS subtypes are autosomal recessive. MPS II is one of the exceptions that inherited in an X-linked manner [Neufeld EF and Muenzer J, 2001]. Most of the patients are males however there are a few cases of well-documented MPS II female patients [Winchester B *et al.*, 1992; Sukegawa K *et al.*, 1998; Cudry S *et al.*, 2000]. For those female patients of mild form, the mutant allele is most likely to occur as heterozygote. Their clinical manifestations are the results of skewed X-chromosome inactivation which prevents the expression of the normal allele.

1.1.1 Prevalence of MPS II

All types of MPS are rare genetic diseases. MPS II is the most prevalent subtype among MPS worldwide [Neufeld EF and Muenzer J, 2001; Kato T *et al.*, 2002]. Patients were identified in the western populations as well as in the Asian populations such as Japanese [Sukegawa K *et al.*, 1995; Isogai K *et al.*, 1998; Kato T *et al.*, 2005], Korean [Kim CH *et al.*, 2003], and Taiwanese [Chang JH *et al.*, 2005]. The average estimated incidence of MPS II is about 1 in 110,000 live newborn males [Tomanin R *et al.*, 2002]. In Northern Ireland, there is 1 in 72,000 male live births inherited MPS II [Nelson J. 1997]. And there is 1 in 165,000 male live births affected in Western Australia [Nelson J *et al.*, 2003].

In Hong Kong, we have different types of LSDs. According to the Joint Metabolic Clinic under Prince of Wales Hospital (data published in 2004), LSDs are relatively more common than the other metabolic disorders (Table 1.1). There are totally 20 patients affected by various types of LSDs, of which only 3 patients are inherited MPS II. We also have three other MPS subtypes, i.e. type I, type III and type VI (Table 1.2). In Hong Kong, MPS III is relatively more common than the other subtypes. Besides the MPS, the following LSDs are also found in Hong Kong, i.e., mucolipidosis type II and Niemann-Pick C.

Table 1.1 Inborn error of metabolism in Hong Kong

Inborn errors of metabolism	No. of patients
Lysosomal storage diseases (LSDs)	20
Disorders of fatty acid oxidation	10
Peroxisomal disorders	10
Mitochondrial disorders	9
Disorders of carbohydrate metabolism	8
Organic acidurias	8
Disorders of transport & mineral metabolism	6
Urea cycle disorders	2

Table 1.1 There are various types of metabolic disorders in Hong Kong. The data was published by the Joint Metabolic Clinic under Prince of Wales Hospital in August 2004.

Table 1.2 Lysosomal storage diseases (LSDs) in Hong Kong

Lysosomal storage diseases	No. of patients (Total = 20)
Mucopolysaccharidosis Type III	5
Mucopolysaccharidosis Type I	3
Mucopolysaccharidosis Type II	3
Mucopolysaccharidosis Type VI	3
Mucolipidosis Type II	3
Niemann Pick Type C	3

Table 1.2 There are mainly three different types of LSDs in Hong Kong. The data was published by the Joint Metabolic Clinic under Prince of Wales Hospital in August 2004.

1.1.2 Pathophysiology of MPS II

The deficiency of the housekeeping enzyme IDS causes lysosomal accumulation of undegraded GAG affecting various tissues and organs such as heart valves, liver, lung and kidney [Neufeld EF and Muenzer J, 2001; Muenzer J *et al.*, 2002; Tomanin R *et al.*, 2002]. For instance, vacuolization of lymph nodes and synoviocytes in the limbs account for swollen joints and hinder the joints movement [Muenzer J *et al.*, 2002]. Depending on the dynamic balance between the level of storage products and residual enzyme activity, the patients will develop various chronic and progressive patterns of clinical severity [Tomanin R *et al.*, 2002; Parkinson EJ *et al.*, 2004]. Indeed, the clinical heterogeneity is the results of different mutations at the *IDS* gene affecting enzyme expression level, protein stability, or catalytic function [Wilson PJ *et al.*, 1990; Hopwood JJ *et al.*, 1993; Tomatsu S *et al.*, 2004].

1.1.3 Clinical features of MPS II

MPS II patients are usually normal at birth. When GAG begins to accumulate abnormally, various organs become dysfunction and symptoms develop [Tomanin R *et al.*, 2002]. As demonstrated in a knockout mouse model of MPS II, swollen joints were visualized from 4 weeks of age. And the vacuolization was widely spread at 60 weeks [Muenzer J *et al.*, 2002]. In human, GAG deposits in the testes and leads to sterile male patients. And the patients develop cardiovascular disorders because of GAG deposition in the intima. Besides, deposits of GAG in the central nervous system may lead to mental retardation. Other clinical phenotypes include hepatosplenomegaly, dysostosis with dwarfism, deafness, coarse facies and massive excretion of heparan sulfate in urine. In general, the manifestations of MPS II are

less severe than MPS I and MPS II does not affect the cornea [Neufeld EF and Muenzer J, 2001].

MPS II patients have a broad spectrum of clinical phenotypes [Neufeld EF and Muenzer J, 2001]. The classification of clinical severity depends on the features including age of onset, survival period, skeletal changes, developmental delay and the presence of progressive mental retardation [Neufeld EF and Muenzer J, 2001]. Mainly, MPS II can be classified into a mild form and a severe neuronopathic form. The mild form involves slow deterioration and fairly normal intelligence while the severe form is early onset and mentally retarded. Some of the patients are regard as intermediate between the two extremes [Neufeld EF and Muenzer J, 2001].

The patients who are mildly affected usually preserve normal intelligence with no mental deterioration. They have obvious somatic changes of joint stiffness and short status. The symptoms are late onset with reduced rate of progressing. The patients can survive into the fourth or fifth decade of life [Young ID *et al.*, 1982; Neufeld EF and Muenzer J, 2001]. In contrast, the severe form has global retardations leading to cardiovascular disorders, respiratory diseases, hepatosplenomegaly and mental retardation. The patients have short stature with stiff joints and skeletal deformities which seriously disturbing their behaviour. Also, the patients have coarse facial features that can be easily recognized since 2- to 4-year-old. Due to the progressive deterioration of many organs and early onset, patients suffering from the debilitating diseases lead to an early demise before the age of fifteen years [Young ID *et al.*, 1982; Neufeld EF and Muenzer J, 2001].

1.1.4 Clinical management of MPS II

1.1.4.1 Diagnostic methods for MPS II

GAG deposits in the kidney leading to massive excretion of GAG in urine and forms the basis for a routine biochemical diagnostic test for MPS. Analysis of urinary lysosomal membrane-associated protein (LAMP-1) may also serve as a marker for MPS. Further quantification of the specific mucopolysaccharides in the urine samples can confirm the diagnosis of most MPS subtypes. However, MPS I and MPS II share the same biochemical markers, heparan sulfate and dermatan sulfate. Therefore, residual enzyme activity should be tested in order to differentiate the two subtypes. MPS I patients have α-L-iduronidase (IDUA) deficiency while MPS II patients show IDS deficiency.

IDS deficiency can be determined in a variety of cells and body fluids. The activity assay can be simply performed using patients' serum or plasma samples. Cultured fibroblasts, developed lymphoblastoid cells, peripheral blood leukocytes and hair roots can also be obtained from the patients for enzyme assays [Li P, Bellows AB and Thompson JN, 1999]. Actually, residual enzyme activity itself does not provide comprehensive information for MPS II diagnosis. Usually, a low level of active enzyme is enough for maintenance of normal phenotypes in most LSDs but the minimum amount of IDS is unknown [Thomas GH, 1994]. Additionally, the IDS protein level and its residual activity show little or no direct correlation with the clinical severity of MPS II [Parkinson EJ *et al.*, 2004; Kato T *et al.*, 2005]. Enzyme assays are also not feasible for carrier detection because the activities of carriers overlap with normal range [Lyon MF, 1961; Cudry S *et al.*, 2000].

RT-PCR sequencing for *IDS* cDNA or direct PCR sequencing for *IDS* genomic DNA is more feasible to characterize the molecular lesions in MPS II patients

[Isogai K *et al.*, 1998; Li P, Bellows AB and Thompson JN, 1999]. DNA sequencing of the *IDS* gene serves as an excellent tool for direct diagnosis of MPS II because it provides definitive results and is automated. Additionlly, it is reliable for carrier detection and prenatal diagnosis. Direct sequencing for the coding sequences and the exon-intron boundaries of the *IDS* gene helps to identify small mutations like single nucleotide changes, deletions or insertions. For detecting gross structural alterations, we can use southern blot analysis. Other approaches like direct restriction enzyme digestion, single-strand conformation polymorphism (SSCP) analysis, heteroduplex analysis with mutation detection enhancement (MDE) gel electrophoresis are also useful for mutational analysis [Isogai K *et al.*, 1998].

1.1.4.2 Treatments for MPS II

Bone marrow transplantation (BMT), enzyme replacement therapy (ERT), substrate deprivation therapy and enzyme enhancement therapy are some of the symptomatic therapies for most LSDs. The allogeneic BMT has been described for treating several MPS subtypes including type II, type IV and type VI with evidences of metabolic correction in various tissues and organs [Warkentin PI *et al.*, 1986; Coppa GV *et al.*, 1995]. However, only MPS II patients with mild form can be considered for this procedure [Coppa GV *et al.*, 1995]. The application of BMT is also limited by the lack of histocompatible donors, its complications such as graft-versus-host disease, and the significant mortality rate up to 20 - 50 % [Warkentin PI *et al.*, 1986].

New therapeutic measures of MPS II are under active developments. Approaches like ERT can ameliorate the signs and symptoms of some LSDs. The first enzyme available for ERT was Alglucerase in treating the Gaucher disease [Whittington R and Goa KL, 1992]. Nowadays, we have two more enzyme products

developed. They were approved in European and American markets for treating Fabry disease and MPS I since the early 2000s. Currently, there are ongoing trials for glycogen storage disease II/ Pompe disease, MPS II and MPS VI [Muenzer J *et al.*, 2002]. ERT for MPS II started the Phase I/II clinical trial since 2002. It showed pronounced effects of GAG reduction in the heart, lung, liver, kidney, spleen, skin and skeletal muscle of the animal models [Muenzer J *et al.*, 2002]. However, purified enzyme may be quite expensive for the patients [Tomanin R *et al.*, 2002].

Gene therapy is another potential treatment for MPS II patients started the Phase I/II clinical trial since early 2000s [Braun SE *et al.*, 1993; Whitley CB *et al.*, 1996; Stroncek DF *et al.*, 1999]. Mainly, amphotropic retroviral and adenoviral containing wild-type human *IDS* coding sequence are considered for transduction due to their engineered non-immunogenic microcapsules. Recent study suggested that MT-based vectors also have great potential strategy of gene therapy utilizing human CD34+ stem cell as targets. [Hong YT *et al.*, 2003] And the recombinant IDS enzyme synthesized in over-expressing cell clones such as macrophagic cells would be therapeutic for GAG clearance in the central nervous system [Braun SE *et al.*, 1993; Krivit W *et al.*, 1995; Tomanin R *et al.*, 2002].

1.2 Iduronate-2-Sulfatase protein (IDS)

1.2.1 Role in GAG degradation

IDS (EC 3.1.6.13) is a house keeping lysosomal enzyme. Expression of IDS is found in all types of cells except mature red blood cells. The enzyme is classified as sulfuric ester hydrolase acting on the sulfated monosaccharide residue. It belongs to a sulfatases family containing at least nine members that hydrolyze sulfate esters in human cells. IDS protein has high degree of homology with the other mammalian sulfatases but each of them displays absolute substrate specificity [Wilson PJ *et al.*, 1990]. Human IDS is responsible for hydrolysis of the C2-sulfate ester bond from non-reducing-terminal L-iduronic acid residues in the heparan sulfate and dermatan sulfate [Neufeld EF and Muenzer J, 2001]. It is the first enzyme catalyzes the breakdown of the storage GAG compounds. The sequential degradation pathway is shown in Figure 1.1. Besides sulfatases, there are glycosidases and nonhydrolytic transferase involved in the pathway.

Figure 1.1 Glycosaminoglycan degradation pathways

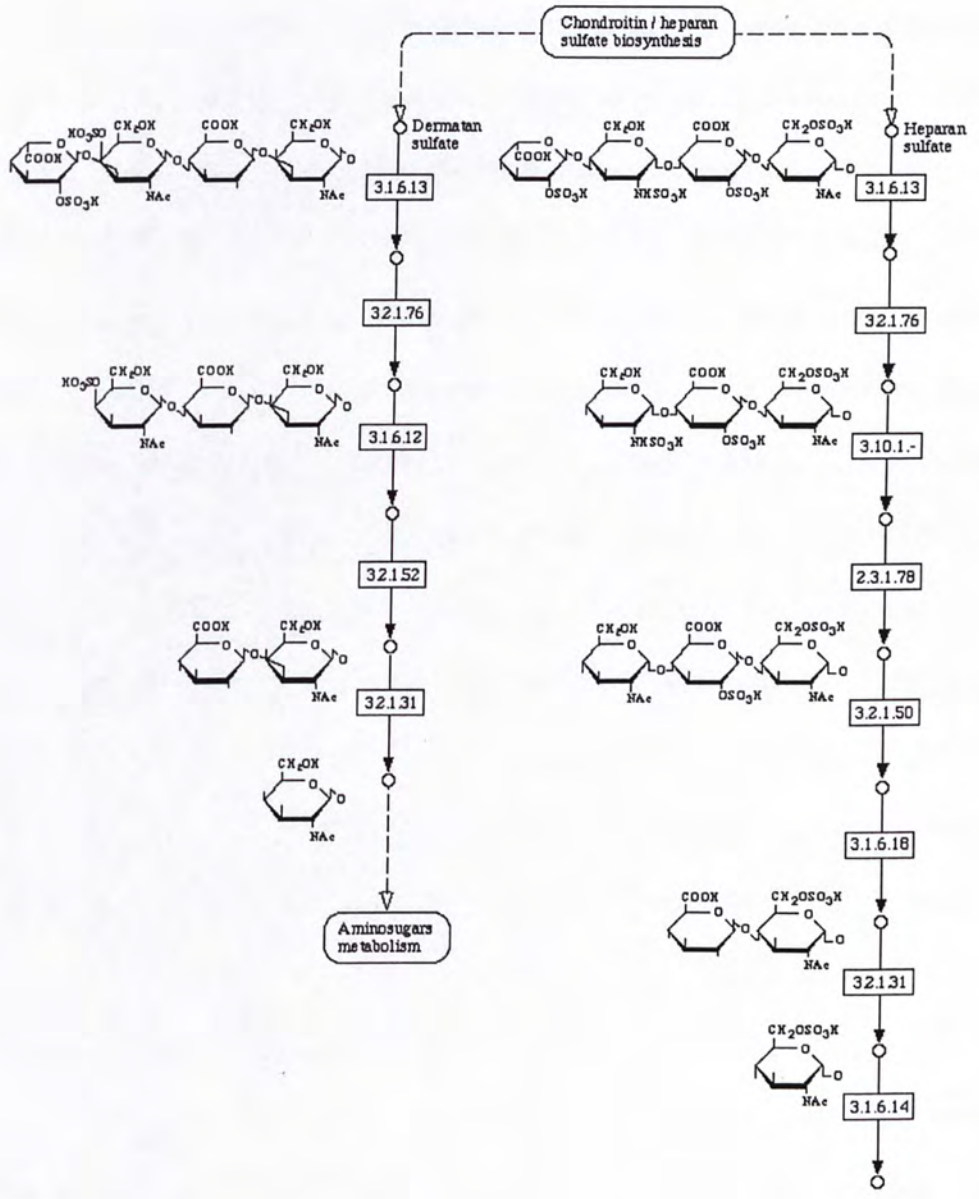


Figure 1.1 The stepwise degradation of heparin sulfates and dermatan sulfatase. The carbohydrate chains of GAG are sequentially removed by a series of lysosomal enzymes. IDS is encoded as EC 3.1.6.13 which initiate the reaction. The entire first step of degradation is completed by another enzyme, IDUA (EC 3.2.1.76). Deficiency of IDS or IDUA results in large amount of herparan sulfase and dermatan sulfase excreted in the urine.

1.2.2 Post-translational modifications

There are eight lysosomal sulfatases including IDS and three non-lysosomal sulfatases in the human cells. Some of them may be evolutionarily related to a common ancestral gene and show similar amino acid sequences (Appendix 1 and 2), structures and catalytic mechanisms [Wilson PJ *et al.*, 1990]. Like many mammalian sulfatases, IDS polypeptide is synthesized as a larger precursor which is composed of 550 amino acids. Then, it undergoes several secondary modifications during the transport from endoplasmic reticulum (ER) to lysosomes (Figure 1.2). Although IDS is known to have no disulphide bonds, there are other post-translational processings include formylglycine (FGly) formation, glycosylation, phosphorylation of oligosaccharide chains, signal peptides cleavage, an N-terminal proteolytic step downstream of residue 33, and a C-terminal proteolytic step upstream of residue 456 [Wilson PJ *et al.*, 1990; Froissart R *et al.*, 1995]. Furthermore, there is an internal cleavage to form additional mature IDS of 45-44 kDa [Millat G *et al.*, 1997a].

1.2.2.1 Formylglycine formation

For catalytic activity, all prokaryotic and eukaryotic sulfatases require a common form of post-translational modification at the active site that converts cysteine into 2-amino-3-oxopropionic acid residue (Cα-formylglycine) [Schmidt B *et al.*, 1995]. The pentapeptide C-X-P-S-R (IDS residues 84 – 88) is highly conserved in all human sulfatases for catalytic activity (Appendix 1). Therefore, functional IDS requires FGly conversion at Cys84 residue [Schmidt B *et al.*, 1995; Millat G *et al.*, 1997a]. ER-residential FGly generating enzyme (FGE) is responsible for such conversion either as co-translational or early post-translational modification prior to folding of the newly synthesized polypeptides [Baenziger JU, 2003;

Preusser-Kunze A *et al.*, 2005]. FGE is encoded by the sulfatase modifying factor 1 (SUMF 1) gene located on human chromosome 3p26. The gene has orthologs in all eukaryotes and most prokaryotes except *E.coli*, *S. cerevisiae* and *C. elegans* [Landgrebe J *et al.*, 2003]. There should be alternative FGly formation mechanisms for their sulfatases.

1.2.2.2 Glycosylation

IDS polypeptide has 8 potential N-linked glycosylation sites (Asn-X-Ser/Thr) at residues Asn31, Asn115, Asn144, Asn246, Asn280, Asn325, Asn513 and Asn537 [Wilson PJ *et al.*, 1990]. Glycosylation takes place in the endoplasmic reticulum and the oligosaccharide chains are then phosphorylated. The first potential site (Asn31 residue) is lost during N-terminal proteolysis. The molecular weight of the C-terminal 18 kDa polypeptides suggested that the two N-glycosylation sites at Asn513 and Asn537 residues are used. However, the total number of glycosylated sites in the mature forms remains unknown [Wilson PJ *et al.*, 1990; Froissart R *et al.*, 1995; Millat G *et al.*, 1997b]. Previous study revealed that none of the potential sites is essential for lysosomal targeting [Tikkanen R *et al.*, 1995; Millat G *et al.*, 1997a]. It is believed that high level of complex-type oligosaccharide chains on circulating IDS has functions to promote protein stability and mask its immune reactivity [Parkinson-Lawrence E *et al.*, 2005].

1.2.2.3 Proteolysis

IDS is synthesized initially as 76 kDa precursor and glycosylated to 90 kDa precursor. They are converted through a 62 kDa intermediate into the mature polypeptides of 55 and 45 kDa [Millat G *et al.*, 1997a]. For N-terminal processing, propeptide of eight amino acids (residues 26 - 33) is removed immediately following

the signal peptides cleavage (residues 1 - 25) shortly after transferred into lysosomes [von Heijne G, 1986]. For the first C-terminal proteolytic cleavage (C-proteolysis 1), removal of 18 kDa polypeptide produces the 62 kDa intermediate form [Froissart R *et al.*, 1995]. Previous study sequenced the 18 kDa polypeptide and found that it contains the 95 amino acids at the C-terminal end of IDS precursor [Wilson PJ *et al.*, 1990]. For the second C-terminal proteolytic cleavage (C-proteolysis 2), a 7-8 kDa polypeptide is released and produces the other mature form of IDS with 45-44 kDa [Millat G *et al.*, 1997a].

Figure 1.2 IDS processing steps from precursors to mature forms

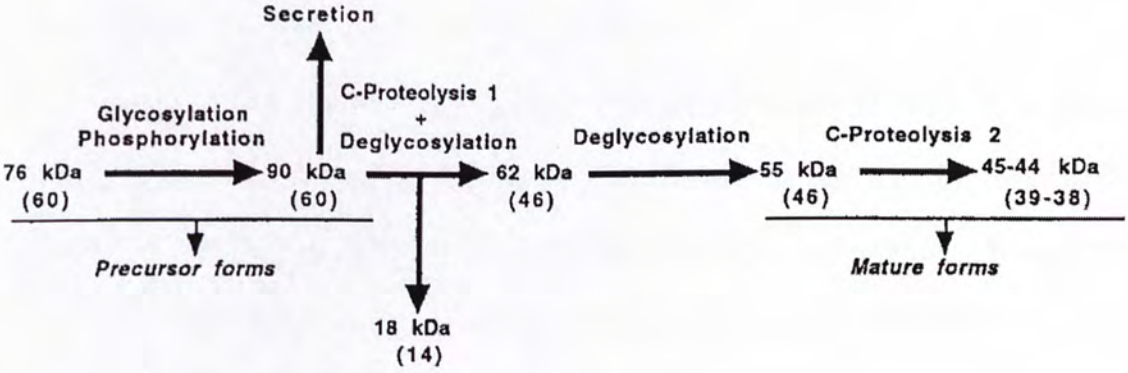


Figure 1.2 IDS processing in COS-7 cells is identical as in human fibroblast. [Millat G *et al.*, 1997b] Precursor IDS processing to mature forms involve FGly formation, glycosylation, N-terminal proteolytic cleavages and C-terminal proteolytic cleavages. Molecular masses of deglycosylated polypeptides are indicated in parentheses (kDa).

1.2.3 Iduronate-2-sulfatase gene (*IDS*)

The *IDS* cDNA containing the complete coding sequence was first isolated in 1990 [Wilson PJ *et al.*, 1990]. The genomic organization and nucleotides sequence of *IDS* gene were then reported in 1993 [Flomen RH *et al.*, 1993; Wilson PJ *et al.*, 1993]. The gene has 1650 base pairs (bp) open reading frame and is mapped to Xq27.3-q28 boundary of the long arm of the X-chromosome spanning approximately 24 kb [Wilson PJ *et al.*, 1990; Flomen RH *et al.*, 1993; Wilson PJ *et al.*, 1993; Timms KM *et al.*, 1995]. The potential promoter for *IDS* gene lacks a TATA box but like many other housekeeping genes, it contains a GC box consensus sequence that span from 400 bp upstream to 150 bp downstream of the initiation codon ATG. The cDNA sequences are numbered in accordance with the A of the ATG initiation codon shown as "+1" (Appendix 3).

There are two transcription variants synthesized from the same *IDS* locus (Figure 1.3). Including the untranslated region, one of the variant (GenBank accession number: NM_000202) is a 2.3 kb transcript containing 9 exons [Wilson PJ *et al.*, 1990]. The alternative form (GenBank accession number: NM_006123) is a 1.4 kb transcript containing only exon 1 to exon 7 with a distinct C-terminal end [Malmgren H *et al.*, 1995]. Both exons 8 and 9 are absent in this alternative variant and the 207 amino acids at the C-terminal end of variant 1 are replaced by 7 amino acids (F-L-M-R-T-N-T). Besides, there is an *IDS*-like pseudogene located about 20 kb distal from the telomeric end of the active *IDS* gene [Timms KM *et al.*, 1995]. This pseudogene comprises the copies of *IDS* exons 2, 3 and intron 7 which have homologous regions with the active gene.

Figure 1.3 Genomic structure of *IDS* gene

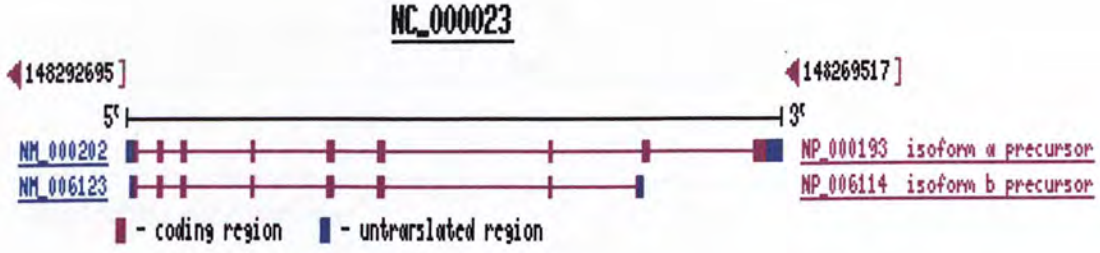


Figure 1.3 Exon/ intron organization of the human *IDS* gene. There are two transcription variants. In this study, the mutations were located at exon 8 and which does not contain in the transcription variant 2. Thus, the *IDS* cDNA was cloned from transcription variant 1, i.e. NM_000202.

1.2.3.1 Properties of *IDS* mutations

Over 300 mutations underlying MPS II were identified worldwide indicating a high degree of allelic heterogeneity at *IDS* locus [Stenson PD *et al.*, 2003; Human Gene Mutation Database]. The wide spectrum of genetic alterations including deletions, insertions, point mutations, splice site mutations, frameshift and homologous recombination imply for the pathogenesis and the observed clinical heterogeneity. Approximately 89 % of mutations are small lesions affecting 1 to 20 bp (Table 1.3). Of which, about 70.4 % are point mutations predicting amino acids replacement (missense), premature termination of translation (nonsense) and aberrant splicing. 13 % of patients have homologous intrachromosomal recombination with *IDS* pseudogene [Bondeson ML *et al.*, 1995]. A loop-cleavage-repair model suggested that homologous regions exchange between *IDS* intron 7 and sequences close to exon 3 of pseudogene [Bondeson ML *et al.*, 1995; Birot AM *et al.*, 1999]. Patients carrying structural rearrangements or gross alterations result in clinically severe MPS II.

Table 1.3 Allelic heterogeneity at *IDS* locus

Mutation type	Total no. of mutations	Percentage (%)
Nucleotide substitutions (missense/ nonsense)	171	53.61
Nucleotide substitutions (splicing)	29	9.09
Small deletions	58	18.18
Small insertions	23	7.21
Small indels	3	0.94
Gross deletions	26	8.15
Gross insertions & duplications	1	0.31
Complex rearrangements (including inversions)	8	2.51
Total	319	100

Table 1.3 Over 300 different mutations underlying *IDS* gene result in mucopolysaccharidosis type II. Mostly, the patients inherited point mutations predicting amino acid replacement and premature termination of translation. Only 11 % of mutations involve gross structural alterations that result in clinically severe MPS II.

1.2.3.2 Methylation patterns are correlated with transitional mutations

The point mutations are nonrandomly distributed over the exons and relatively more frequent in exons 3, 5, 8 and 9 of *IDS* gene (Figure 1.4 & Table 1.4) [Rathmann M *et al.*, 1996]. They are likely to correlate with methylation patterns of CpG sites in the exons [Tomatsu S *et al.*, 2004]. And almost all recurrent transitional mutations are associated with CpG sites. For instance, codons Ala85, Arg88, Ser333, Arg443 and Arg468 with CpG dinucleotide hypomethylated, are the mutational hotspots [Isogai K *et al.*, 1998; Li P, Bellows AB and Thompson JN, 1999]. Over 35 % point mutations found in CpG sites and nearly 80 % are derived from C>T or G>A transitional events by in situ methylation-deamination process [Cooper DN and Krawczak M, 1993; Rathmann M *et al.*, 1996; Li P, Bellows AB and Thompson JN, 1999; Tomatsu S *et al.*, 2004].

For mammalian housekeeping genes, CpG islands are normally located in the 5′ region and mostly unmethylated. And the general patterns of methylation throughout the *IDS* active allele (Xa) are very similar between normal females and normal males without any difference between the sexes [Tomatsu S *et al.*, 2004]. However, the *IDS* inactive allele (Xi) has methylation patterns quite different from Xa because of its nature of X-chromosomal inactivation among the females. All the CpG sites on Xi are characteristically hypermethylated and highly mutable [Gartler SM and Riggs AD, 1983; Park JG and Chapman VM, 1994]. The asymmetrical methylation patterns of both *IDS* alleles in the female carriers may lead to symptomatic MPS II when non-random inactivation of X-chromosome occurs [Cudry S *et al.*, 2000]. Nevertheless, de novo mutations tend to occur frequently in male meioses [Rathmann M *et al.*, 1996; Froissart R *et al.*, 1998].

Figure 1.4 Mutation map of *IDS* gene

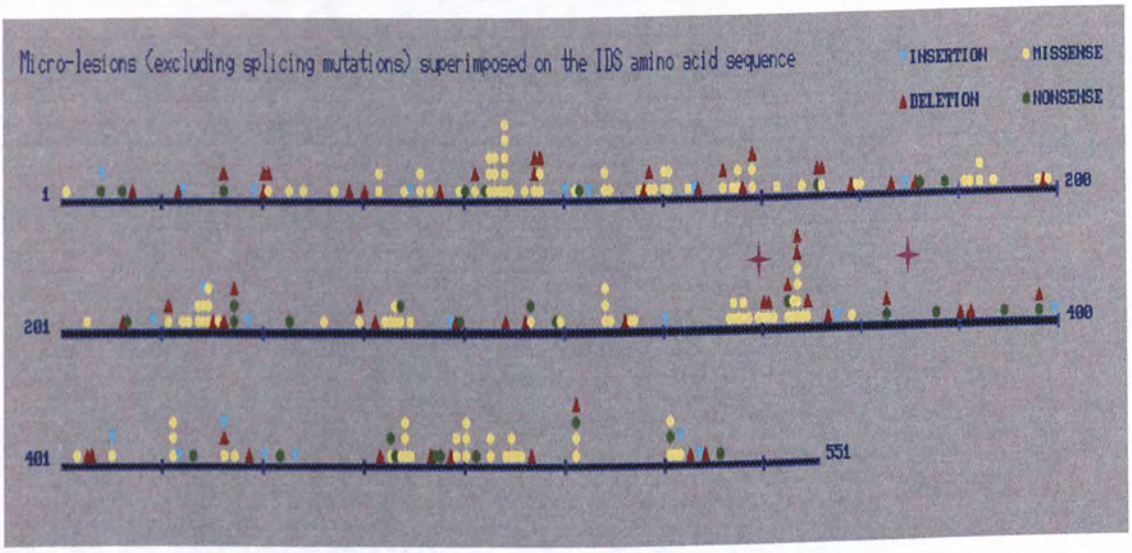


Figure 1.4 The mutation map was obtained from the Human Gene Mutation Database. There are 171 point mutations reported on *IDS* gene responsible for MPS II. Exons 3, 5, 8 and 9 are the hot spots of missense mutation. Nearly 25 % of point mutations were located in exon 9. Previously unreported mutations are indicated as star on the map, i.e. p.S369X and p.L339P.

Table 1.4 Non-random distribution of point mutations on *IDS* gene

Exon number	1	2	3	4	5	6	7	8	9
Frequency (total = 171)	4	12	38	6	24	13	11	21	42

Table 1.4 The data of mutation frequency was obtained from the Human Gene Mutation Database. In this study, two novel mutations were identified at exon 8 but they were not included in the table.

1.2.3.3 Genotype-phenotype correlations between *IDS* gene and MPS II

Several studies showed presence of genotype-phenotype correlations between *IDS* gene mutations and MPS II by structural analysis, or more comprehensively by *in vitro* expression [Rathmann M *et al.*, 1996; Millat G *et al.*, 1997a; Villani GRD *et al.*, 2000; Bonuccelli G *et al.*, 2001; Ricci V *et al.*, 2003; Chang JH *et al.*, 2005; Kato T *et al.*, 2005]. Most small deletions from 1 to 3 bp lead to frameshift and premature chain termination. For those mutations predicting termination of translation located very near to the C-terminal end, e.g. Q531X may cause a mild phenotype due to the normal C-terminal proteolytic processing [Bunge S *et al.*, 1992]. However, the premature stop codon happened at other regions may affect protein synthesis and/ or stability. For instance, W345X and 787delG (A263X) resulted in unstable truncated enzyme protein and cause clinically severe MPS II [Bunge S *et al.*, 1992; Kim CH *et al.*, 2003].

Genetics alterations belong to the conservative changes at sulfatase C-X-P-S-R motif are correlated with severe forms of MPS II. The mutations at Cys84 residue have drastic effect for IDS processing and abolish the enzymatic activity. [Millat G *et al.*, 1997a] IDS Arg88 residue is not essential for processing but contributes to normal transport of the polypeptides from the early vacuolar compartment to lysosomes [Villani GRD *et al.*, 2000]. Substitution studies of Pro86 residue also revealed failure of lysosomal targeting and lead to severe disease [Millat G *et al.*, 1998]. Mutations at other residues, e.g. W267C and S117del also lead to partial targeting or no lysosomal targeting of the IDS polypeptides resulting in severe phenotypes [Rathmann M *et al.*, 1996; Chang JH *et al.*, 2005]. Mistargeted IDS polypeptides may be inactive due to the loss of lysosomal environment which provides optimum pH for catalytic activity.

Some mutations, such as S333L and G336V may change the active site

conformation at adjacent residues Asp334 and His335 cause severe phenotypes by reducing the enzymatic activity [Isogai K *et al.*, 1998; Froissart R *et al.*, 1998; Kato M *et al.*, 2005]. Mutations like K347T, N265I and 473delTcc may also have possible effects on enzyme conformation and catalytic function [Bonuccelli G *et al.*, 2001]. For those mutations not conserved for IDS processing or lysosomal targeting or not structurally important for functional activity may cause mild to intermediate phenotypes. For instance, the amino acids alterations at Arg95 residue changed to lysine, leucine or threonine are responsible for intermediate phenotype [Dierks T *et al.*, 1999; Goldenfum SL *et al.*, 1996; Moreira da Silva I *et al.*, 2001].

1.3 In this study

To date, there are totally 4 individuals affected by MPS II in Hong Kong. As mentioned in Chapter 1.1, mutations of the *IDS* gene are responsible for the disease. However, no mutation data and family pedigree for the patients was published before. In this study, there are three main objectives: 1) to identify the genetic defects lying in the *IDS* gene for all the MPS II patients in Hong Kong; 2) to examine the deleterious effects of the mutations so that possible correlations between the mutation data and clinical phenotypes can be estimated, and; 3) to investigate the mechanism of post-translational processing for IDS polypeptides maturation.

1.3.1 Mutational analysis

To my knowledge, this is the first time to study the molecular lesions for MPS II patients in Hong Kong. In this study, mutations of the *IDS* gene were identified comprehensively by direct sequencing the entire coding sequence. Restriction analysis was carried out to confirm the absence of the novel mutation in 50 normal individuals. The identified mutations were used for prenatal diagnosis and carrier detection among the patients' families.

1.3.2 *In vitro* expression of mutant IDS

The major evidence of causality for novel mutations should preferably involve well-designed expression studies of the mutant enzyme [Wilson PJ *et al.*, 1990; Daniele A *et al.*, 1993; Bonuccelli G *et al.*, 2001]. The mutant cDNA was synthesized by site-directed mutagenesis which was commonly used for IDS studies since the early 1990s [Annella T *et al.*, 1993]. Two different *in vitro* expression systems were used for evaluating the functional consequences of the novel missense mutation at codon Leu339. Besides the pathological effects, the structural importance of this residue was also verified by substitution studies.

Three complementary experiments were involved in this study. Although human cell lines allow better processing of IDS precursors, COS-7 cells was proved to have identical processing steps of IDS precursors as in human fibroblasts [Millat G *et al.*, 1997b]. The IDS mutants were expressed in COS-7 cells for 1) investigating IDS enzymatic defect by fluorometric activity measurement; 2) investigating *IDS* mRNA stability by RT-PCR study. The IDS mutants were also expressed in cell-free *in vitro* expression systems for 3) examining IDS precursors stability by simple protein staining. The cell-free expression systems are coupled with transcription and translation in a single reaction tube for recombinant protein production. This alternative approach becomes one of the most important tasks in biotechnology today.

This is the first time to apply cell-free expression systems to synthesize IDS protein. The performance of microbial systems, plant systems and mammalian systems were compared (Table 2.4). It is believed that both the plant and mammalian systems provide eukaryotic environments for eukaryotic mRNA binding and facilitate the translation process. Simultaneously, microbial expression systems

continue to be wildly used because of high yields and ease of manipulations [Betton JM, 2003; Voloshin AM and Swartz JR, 2005; Graumann K and Premstaller A, 2006]. However, one of the largest problems is the formation of inclusion bodies in microbial system. When large amount of foreign proteins are synthesized quickly, the polypeptides will be misfolded and aggregated. The cell-free systems have the advantages of being open and flexible, supplements can be added to assist protein folding or even processing in this study.

1.3.3 Maturation of IDS polypeptides

Previous studies revealed that none of the potential N-glycosylation sites is essential for lysosomal targeting although the mature IDS polypeptides have molecular weight suggesting that they are glycosylated [Millat G *et al.*, 1997a]. In this study, the importance of glycosylation during IDS processing was studied by expressing wild-type IDS in the cell-free systems which do not allow glycosylation takes place. And the unglycosylated proteins were introduced into COS-7 cells for verifying the cellular uptake mechanism of IDS. Additionally, previous studies suggested that IDS processing can be pre-lysosomal or completed just after the polypeptides entered the lysosomes [Schmidt B *et al.*, 1995]. In this study, a modified *IDS* cDNA without the signal peptide was constructed for estimating the location of post-translational processing.

Chapter 2 Materials and Methods

2.1 Mutation screening for MPS II patients

2.1.1 Patients

Four male patients were diagnosed to be MPS II by clinical observations and biochemical testing previously. They had significant hyper-excretion of urinary heparin sulfate and dermatan sulfate detected by electrophoresis. Their genetic defects lying in *IDS* gene were investigated in this study. Patient 1 was diagnosed at 14-year-old with severe phenotypes. Patient 2 and Patient 3 are brothers preserved normal intelligence and affected mildly by the disease. The screening for *IDS* mutations was done at their age of 30s. Patient 4 was identified four years after birth and was affected severely by the disease. In this study, the female relatives of the patients were also subjected to carrier detection by mutational analysis.

2.1.2 Genomic DNA extraction

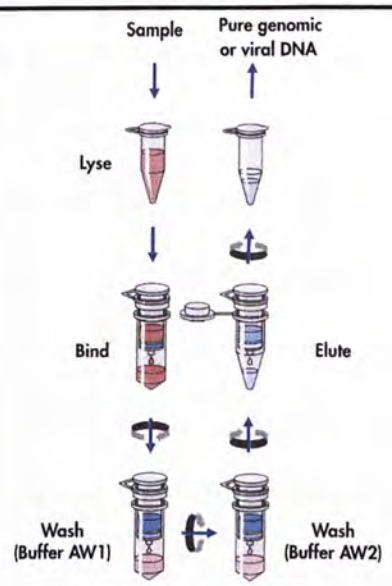
2.1.2.1 Materials

Peripheral blood was obtained from the four MPS II male patients and six of their family members including five female relatives. Their genomic DNA was extracted from the whole blood samples using the commercial kit described below according to the manufacturer's instruction.

QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) comprises protease solvent, QIAGEN protease, buffer AL, buffer ATL, buffer AW1, buffer AW2, buffer AE, QIAamp spin columns, 2 mL collection tubes and 1.5 mL centrifuge tubes. QIAGEN protease is supplied lyophilized and required to be dissolved by the protease solvent before use. All reagents are stored at room temperature (RT) except the dissolved protease which is stored at -20°C. The protocol claims to have a yield of up to12 µg DNA from 200 µL whole human blood. The purified DNA is from 200 bp up to 50 kb depending on the age and storage of blood samples, with majority of 20-30 kb.

Ethanol absolute GR for analysis (Merck KGaA, Darmstadt, Germany) C_2H_5OH, with MW 46.07, is stored at RT. It is highly flammable and should be stored in fume cupboard.

Figure 2.1 QIAamp DNA Blood Mini Kit



Adapted from Qiagen, 2004

QIAamp DNA Blood Mini Kit isolates DNA from blood and other body fluids. DNA is released from lysed samples and binds specifically to the silica-gel membrane in the QIAamp spin column during centrifugation. Salt and pH conditions are adjusted to allow optimal binding of DNA onto the membrane, while contaminants pass through. Any residual contaminants which can act as PCR inhibitors, such as divalent cations and proteins are completely removed in two subsequent wash steps, leaving pure DNA ready to be eluted.

2.1.2.2 Methods

The process was summarized in Figure 2.1. A mixture of cell lysis reactions was prepared. 200 µL whole blood samples and 20 µL protease were added to 200 µL buffer AL. The mixture was

vortexed for 15 sec and incubated at 56 °C in a heating block for 10 min. 200 µL absolute ethanol was added and mixed by vortex for 15 sec. The mixture was then applied to the QIAamp Spin Column contained in 2 mL collection tube. The filtrated cell debris was discarded after centrifugation at 6000 × g for 1 min. 500 µL buffer AW1 was added, and the column was centrifuged at 6000 × g for 1 min. 500 µL buffer AW2 was added, and the column was centrifuged at 16100 × g for 3 min. The column was further centrifuged at 16100 × g for 1 min to remove any residual buffer AW2. 50 µL buffer AE was then added to the column for 5 min incubation at RT. Finally, DNA was eluted by centrifugation at 6000 × g for 1 min and collected in a clean 1.5 mL centrifuge tube. The DNA yield and purity were determined spectrophotometrically with a BioPhotometer (Eppendorf AG, Hamburg, Germany) by measuring absorbance at 260nm and calculating the A_{260}/A_{280} ratio respectively.

2.1.3 *IDS* exons amplification by Polymerase Chain Reaction (PCR)

2.1.3.1 Materials

2.1.3.1.1 PCR

100 mM dNTP Set (Invitrogen Corporation, Carlsbad, California, USA) is supplied in a concentration of 100 mM and required to be stored at -20°C. It is diluted to 10 mM with dH_2O.

AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, California, USA) is supplied together with 10 × PCR Buffer and $MgCl_2$. The enzyme is stored at -20°C while the other components are stored at 4°C. AmpliTaq Gold DNA Polymerase is a chemically modified form of AmpliTaq DNA Polymerase, with a chemical moiety attached to the enzyme. When the chemical moiety is attached, the

enzyme is inactive. By a heat activation step, the chemical moiety is cleaved and active enzyme is released. Thus, during reaction set-up and the first ramp of thermal cycling, the enzyme is still inactive and mis-primed primers cannot be extended.

Flanking primers (QIAGEN, Hilden, Germany) is customized from QIAGEN. Nine pairs of *IDS* primers, namely IDSexon1F to IDSexon9F and IDSexon1R to IDSexon9R (Table 2.1), were designed using a web-based software program, Primer3. PCR fragments amplified from the corresponding primer pairs include the particular exons and the flanking introns of the *IDS* gene.

2.1.3.1.2 Agarose gel electrophoresis

1 kb DNA ladder (Invitrogen Corporation, Carlsbad, California, USA) is supplied in a concentration of 1 μg/ μL. It is diluted to 100 ng/ μL by loading buffer and TE buffer. It is stored at 4°C. Detailed information is shown in Appendix 4.1.

Ethidium Bromide (Invitrogen Corporation, Carlsbad, California, USA) is supplied in a concentration of 10 mg/mL. It is stored at RT. It is carcinogenic.

Biowest agarose, regular (Gene Company Limited, Hong Kong, China) is manufactured in Europe. It is stored at RT.

Tris-Acetate-EDTA (TAE) electrophoresis buffer (Invitrogen Corporation, Carlsbad, California, USA) is supplied in 10 × solution. It is diluted to 1 × with dH_2O. It is stored at RT.

2% agarose gel was prepared by dissolving 2.4 g Biowest agarose in 120 mL 1 × TAE electrophoresis buffer and heating in microwave oven. The solution was cooled under running tap water to 55°C. 60 μg ethidium bromide was added. The molten mixture was mold in a gel cast with suitable combs inserted. It was allowed to stand at RT until set.

Table 2.1 Primers for screening *IDS* mutations

Primer name	sequence (5' → 3')	Optimized annealing temperature	Products size (bp)
IDSexon1F	ccg Acg Agg Agg TcT cTg T	62 °C	348
IDSexon1R	ggg AAc gAA TgA Tgg ATg AA		
IDSexon2F	Agg AcT cAg gcT Tcc Tcc Tc	60 °C	428
IDSexon2R	AAc AAg ATg Tcc cgc AcA AT		
IDSexon3F	Tgg TTT gAg cTc Tgc ATg Ac	60 °C	408
IDSexon3R	AgA ggg cTc Tgc AAA gAc Ag		
IDSexon4F	gTT ccA cTT gcc cAT TTg TT	56 °C	375
IDSexon4R	AAT gAA gcc AcT gcT ccT gT		
IDSexon5F	Tgc cTg gAA AAc AAg AAA cA	56 °C	469
IDSexon5R	gAT gTA gcc Acc TTc ccT gT		
IDSexon6F	AAA Tgc ATc ccA ggc TTA gA	60 °C	464
IDSexon6R	gcA AcA cTg ccT gTg Tcc TA		
IDSexon7F	gcT gTg AcT cTg Tgg gTg AA	60 °C	454
IDSexon7R	ccA ggA Tcc cAc TTT gTT Tg		
IDSexon8F	ccT gAT TTT gAA TAA Agc Agc A	60 °C	457
IDSexon8R	ccc cAA Agc cTA TgA TTc AA		
IDSexon9F	cAT ATg gAg ccc AgA cAg gT	64 °C	641
IDSexon9R	cgA ccA gcT cTA AcT ccT ccT		

Table 2.1 The flanking primers are used for amplifying the *IDS* exons and also used as sequencing primers to screen mutations.

2.1.3.1.3 PCR fragments purification

MicroSpin S-400 HR columns (Amersham Biosciences, Uppsala, Sweden) are pre-packed with Sephacryl resin and equilibrated in TE buffer. They are stored at 4°C. MicroSpin S-400 HR columns utilize spin-colomn chromatorgrapy to purify DNA for a wide range of applications, such as purification of PCR products from unincorporated primers and nucleotides (dNTPs). Spin-column chromatorgrapy enables desired product passing through the gel while retaining smaller impurities. The resin was resuspended by vortex and the cap was loosened one-fourth turn followed by snapping off the bottom closure before use. Excess TE buffer is then removal by centrifugation at 735 × g for 1 min.

2.1.3.2 Methods

2.1.3.2.1 Amplifying *IDS* exons by PCR

All the 9 exons and their flanking introns of *IDS* gene were amplified by PCR using the genomic DNA extracted from peripheral blood as template. The reactions were carried out in a MJ PTC-200 (MJ Research, Waetham, Incorporated, Massachusetts, USA), with the following parameters: one cycle of Taq polymerase activation at 94 °C for 12 min; 40 cycles of DNA denaturation at 94 °C for 30 sec, annealing at 56 – 64 °C for 45 sec, and extension at 72 °C for 45 sec; one cycle of final extension at 72 °C for 7 min. The optimized annealing temperatures for each pairs of primers are summarized in Table 2.1. The reaction mixture of final volume 25 μL contained 50 ng DNA template, 1 × PCR buffer, 50 nmol $MgCl_2$, 5 pmol dNTPs, 12.5 pmol of forward / reverse primers, and 0.625 U AmpliTaq Gold DNA polymerase. For exon 1, 5% of DMSO was added and the amount of $MgCl_2$ was increased to 62.5 nmol. The PCR products were then analyzed by 2% agarose gel electrophoresis to ensure that the specific bands were in correct sizes.

2.1.3.2.2 Purifying PCR fragments

After removal of storage solution in the MicroSpin S-400 HR column, the PCR products were applied to the top-center of the resin. Purified PCR fragments were collected by centrifugation at 735 × g for 2 min.

2.1.4 DNA sequencing for detecting *IDS* mutations

2.1.4.1 Materials

BigDye Terminator v1.1 Cycle Sequencing Kit (Amersham Biosciences, Uppsala, Sweden) consists of BigDye Terminator v1.1 cycle sequencing mix and 5 × BigDye v1.1/3.1 sequencing buffer. The kit is stored at -20 °C.

Sequencing primers (QIAGEN, Hilden, Germany) is the 25-fold dilution of the flanking primers.

AutoSeq G-50 (Amersham Biosciences, Uppsala, Sweden) are MicroSpin columns prepacked with Sephadex G-50 DNA Grade F, equilibrated in water containing 0.05% Kathon CG/ICP biocide as a preservative. They are stored at RT.

MegaBASE Loading Solution (Amersham Biosciences, Uppsala, Sweden) is stored at -20°C.

2.1.4.2 Methods

2.1.4.2.1 Sequencing reaction

The purified PCR fragments of the 9 *IDS* exons were sequenced to screen for mutations by dideoxynucleotide method [Sanger *et al*, 1977]. The reaction mixture of 10 μL final volume contains 2 μL purified PCR products, 1 pmol forward primers, 1 μL BigDye Terminator v1.1 cycle sequencing mix and 3 μL of 2.5 × BigDye

v1.1/3.1 Sequencing Buffer. Cycle sequencing reaction was carried out in a GeneAmp PCR System 9700 Cycler (Applied Biosystems, Foster City, California, USA), with the following parameters: one cycle of 96 °C for 1 min; 25 cycles of 96 °C for 10 sec, 50 °C for 5 sec, 60 °C for 4 min. The products are stored at -20 °C before further processing. If mutation is found, the results will be confirmed by means of DNA sequencing with the reverse primer.

2.1.4.2.2 Purifying sequencing products

Excess solution was removal by centrifugation at 2000 × g for 1 min. The sequencing reaction products were purified through the G-50 spin column centrifuged at 2000 × g for 1 min to remove the residual dye-labeled dideoxynucleotide terminators.

2.1.4.2.3 Analyzing sequencing results

The purified sequencing products were mixed with 12 μL MegaBase loading solution and heated at a 95 °C heat block for 5 mins. The mixture was sequenced and analyzed automatically by the sequencer ABI PRISM 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, California, USA). The mutation site was identified manually and using BLAST programme online.

2.1.5 Fragment restriction endonuclease analysis

2.1.5.1 Materials

DNA template was extracted from 50 normal control females as described in Chapter 2.1.2.2.

PCR reagents are the same as described in Chapter 2.1.3.1.1 and Chapter 2.1.3.1.2.

Flanking primers (QIAGEN, Hilden, Germany) is customized from QIAGEN. IDS-InterF2 and IDS-InterR3 are used to amplify PCR fragment containing codon 339. The sequences of primers are shown in Table 2.2.

ScrF I, 10,000 U/ mL (New England Biolabs, Ipswich, UK) is supplied with 10 × Buffer 4. The reagents are stored at -20 °C. The enzyme recognizes DNA sequence 5′-CCNGG-′3 and forms sticky ends after cleavage.

2.1.5.2 Methods

The fragments of *IDS* gene containing codon 339 were amplified by PCR from the 50 control females. The reactions were carried out in a MJ PTC-200 (MJ Research, Waetham, Incorporated, Massachusetts, USA), with the following parameters: one cycle of Taq polymerase activation at 94 °C for 12 min; 30 cycles of DNA denaturation at 94 °C for 30 sec, annealing at 60 °C for 20 sec, and extension at 72 °C for 15 sec; one cycle of final extension at 72 °C for 7 min. The reaction mixture of final volume 25 μL contained 50 ng DNA template, 1 × PCR buffer, 50 nmol $MgCl_2$, 5 pmol dNTPs, 12.5 pmol of forward / reverse primers, and 0.625 U AmpliTaq Gold DNA polymerase. The PCR products were then analyzed by 2% agarose gel electrophoresis to ensure that the specific bands were in correct sizes. Then 6 μL PCR products were digested by 5 U *ScrF I* with 1 × buffer 4 in 10 μL final volume. The reaction was performed at 37 °C for 16 hours and terminated at 65 °C for 20 min. The digested PCR fragments were analyzed by 3 % agarose gel electrophoresis.

Table 2.2 Primers for *IDS* cloning

Primer name	Purpose	Sequence (5' → 3')	Grade
IDS-c-Fp	Isolation from total cDNA	cTgTgTTgcgcAgTcTTcATgg	HPLC
IDS-E-Fp	Cloning WT into entry vector	cAccATgccgccAccccggAccggcc	HPLC
IDSt-E-Fp	Cloning N-truncated form into entry vector	cAccATgAcAgATgcTcTgAAcgTT	HPLC
IDS-c-Rp	Isolation from total cDNA Cloning into entry vector	ATggTTggcAAAAcTcAAggcA	HPLC
IDS-Nde-Fp	Cloning WT into pIVEX1.3	ATAATTATcATATgccgccAccccggAccggccg	HPLC
IDSt-Nde-Fp	Cloning N-truncated form into pIVEX1.3	ATAATTATcATATgAcAgATgcTcTgAAcgTTc	HPLC
IDS-Xho-Rp	Cloning into pIVEX1.3	AATTATcTcgAgAggcATcAAcAAcTggAAAA	HPLC
IDS-339P -sen	Site-directed mutagenesis	TcATgggTgggcT**ccA**ggTgAAcATggAg	HPLC
IDS-339P -anti	Site-directed mutagenesis	cTccATgTTcAcc**Tgg**AgcccAcccATgA	HPLC
IDS-339R -sen	Site-directed mutagenesis	TcATgggTgggcT**cgA**ggTgAAcATggAg	HPLC
IDS-339R -anti	Site-directed mutagenesis	cTccATgTTcAcc**Tcg**AgcccAcccATgA	HPLC
IDS-InterR1	Sequencing from +296 bp	gTgTcAggTcTccTgccAgTgA	salt-out
IDS-InterF1	Sequencing from +217 bp	cTcTTccAgAATgccTTTgcgc	salt-out
IDS-InterR2	Sequencing from +859 bp	gAccATAcggcAcAcTgATgTT	salt-out
IDS-InterF2	Sequencing from +760 bp	gAggTcccTgATggccTAcccc	salt-out
IDS-InterF2b	Sequencing from +1152 bp	TgATTccgccTcAcAgTTgATg	salt-out
IDS-InterR3	Sequencing from +1381 bp	TcAgTTcAcggggATTAccAgg	salt-out
IDS-InterF3	Sequencing from +1319 bp	AgcATTTTcgATTccgTgAcTT	salt-out

Note: WT means wild-type *IDS* cDNA; N-truncated means *IDS* cDNA without the first 99 bp. The underlined sequences represent the restriction enzymes recognition sites. The altered codon sequences for site-directed mutagenesis are bolded.

2.2 Isolation of *IDS* cDNA from peripheral blood

2.2.1 Materials

2.2.1.1 Total RNA extraction

Peripheral blood was obtained from a normal female individual with plasma IDS activity 343.1 nmol/4h/mL. The total RNA was extracted from the whole blood samples using the commercial kit described below according to the manufacturer's instruction.

RNeasy Mini Kit (Qiagen, Hilden, Germany) comprises beta-mercaptoethanol, buffer RLT, buffer RW1, buffer RW1, buffer RPE, RNeasy Mini spin columns, 2 mL collection tubes and 1.5 mL centrifuge tubes. 1.5 mL buffer RLT is required to mix with 15 µL beta-mercap before use. All the reagents are stored at RT. The kit claims to have a binding capacity of up to100 µg of total RNA with enriched mRNA longer than 200 bp. Most RNAs smaller than 200 bp, e.g. 5.8S rRNA, 5S rRNA and tRNAs, are selectively excluded.

RNase-free DNase set (Qiagen, Hilden, Germany) comprises *DNase I* and buffer RDD. *DNase I* is supplied lyophilized. The set is stored at 4 °C.

Ficoll-Paque Plus (Sigma-Aldrich, St. Louis, Missouri, USA) is supplied as solution of Ficoll 400 and sodium diatrizoate with a density of 1.077 ± 0.001 g/ mL. The solution is sterile for *in vitro* isolation of lymphocytes and is stored at 4 °C.

Phosphate Buffered Saline/ PBS (Sigma-Aldrich, St. Louis, Missouri, USA) contains NaCl (120 mmol/L), KCl (2.7 mmol/L) and phosphate buffer (10 mmol/L), pH 7.4. It is provided in powder form and stored at RT. To prepare PBS solution, the powder was dissolved to 1 L deionized water (dH_2O) and autoclaved at 121 °C for 15 min.

Ethanol absolute GR for analysis (Merck KGaA, Darmstadt, Germany) C_2H_5OH, with MW 46.07, is stored at RT. It is highly flammable and should be stored in fume cupboard.

2.2.1.2 Reverse-transcriptase PCR (RT-PCR)

Total RNA was extracted from peripheral blood cells as described in Chapter 2.2.2.1.

Oligo(dT)$_{20}$ primer (Invitrogen Corporation, Carlsbad, California, USA) is supplied as 50 µM in dH_2O and stored at -20 °C. It is a string of 20 deoxythymidylic acid residues that hybridized to the poly(A) tail of mRNA and can be used as a primer for first-strand cDNA synthesis with SuperScript III Reverse Transcriptase.

dNTP Mix (Invitrogen Corporation, Carlsbad, California, USA) is supplied as 2.5 mM in highly purified H_2O and stored at -20 °C. It contains 2.5 mM each of dATP, dCTP, dGTP and dTTP.

SuperScript III Reverse Transcriptase, 200 U/µL (Invitrogen Corporation, Carlsbad, California, USA) is supplied as a kit containing 0.1 M DTT and 5 × First-strand buffer. All the reagents are stored at -20 °C. The kit claims to have a yield of cDNA from 100 bp to > 12 kb.

Ribonuclease H/ RNase H, 2U/ µL (Invitrogen Corporation, Carlsbad, California, USA) is stored at -20 °C. It is used for removing mRNA during second strand cDNA synthesis.

2.2.1.3 PCR for amplifying *IDS* cDNA

PCR reagents are the same as described in Chapter 2.1.3.1.1 and Chapter 2.1.3.1.2.

Luciferase T7 control vector (Promega Corporation, Madison, USA) is

supplied in a concentration of 0.5 mg/ mL. The vector contains firefly *luciferase* cDNA.

Flanking primers for IDS cDNA (QIAGEN, Hilden, Germany) is customized from QIAGEN. The five sets of primer pairs (Table 2.2), **1)** IDS-c-Fp and IDS-c-Rp, **2)** IDS-E-Fp and IDS-c-Rp, **3)** IDSt-E-Fp and IDS-c-Rp, **4)** IDS-Nde-Fp and IDS-Xho-Rp, **5)** IDSt-Nde-Fp and IDS-Xho-Rp were designed using a web-based software program, OligoAnalyzer. PCR fragments amplified by the first set of primers are the full length *IDS* cDNA with part of the regulatory sequences. PCR fragments amplified by the second set of primers are the full length wild-type *IDS* cDNA started from ATG to stop codon for TOPO cloning. PCR fragments amplified by the third set of primers are the N-terminally truncated form of *IDS* cDNA with the absence of 99 bp for TOPO cloning. Restriction sites were incorporated into the forth and fifth sets of PCR primers for standard sticky-end cloning. PCR fragments amplified by the forth set of primers are the full length *IDS* cDNA without stop codon for cloning into pIVEX-WG-1.3. PCR fragments amplified by the fifth set of primers are the N-terminally truncated form of *IDS* cDNA also for cloning into pIVEX-WG-1.3.

Flanking primers for luciferase gene (QIAGEN, Hilden, Germany) is customized from QIAGEN. The primer pairs were designed using a web-based software program, OligoAnalyzer (Table 2.3).

Table 2.3 Primers for *Luciferase* cloning

Primer name	Purpose	Sequence (5' → 3')	Grade
Luci-E-Fp	Cloning WT into entry Vector	cAccATggAAgAcgccAAAAA cATA	HPLC
Luci-E-Rp	Cloning WT into entry Vector	cAATTTggAcTTTccgcccTTc	HPLC
Luci-InterR1	Sequencing from +293 bp	ATAAATAAcgcgcccAAcAc	salt-out
Luci-InterR2	Sequencing from +708 bp	cgcAgTATccggAATgATTTg	salt-out
Luci-InterR3	Sequencing from +1131 bp	ATccAgATccAcAAccTTcg	salt-out
Luci-InterF1	Sequencing from +1043 bp	cTATTcTgATTAcAcccgAgg	salt-out
Luci-InterF2	Sequencing from +1377 bp	AcAAcAccccAAcATcTTcgA	salt-out

Note: WT means wild-type *luciferase* cDNA

2.2.2 Methods

2.2.2.1 Extracting total RNA by QIAamp RNeasy Mini Kit

Peripheral blood sample was obtained from a normal individual. 6 mL whole blood samples were added to 6 mL Ficoll solution. The mixture was centrifuged at 2500 rpm for 15 min. The ring of white matter containing peripheral white blood cells (WBC) was isolated and rinsed twice with 13 mL cold PBS. The resuspended WBC was centrifuged at 1500 rpm for 5 min. 600 μL prepared buffer RLT was added and mixed with the pellet. The mixture was then homogenized through the QIAShredder Spin Column by centrifugation at 13,000 rpm for 2 min. 600μL of 70 % ethanol was mixed with the lysate. And the mixture was applied to the RNeasy mini column. The cell debris was filtered by centrifugation at >10,000 rpm for 15 sec and discarded. 350 μL buffer RW1 was added, and the column was centrifuged

at >10,000 rpm for 15 sec. A mixture of 10 μL *DNase I* and 70 μL buffer RDD was added and allowed to stand at RT for 15 min. 350 μL buffer RW1 was added, and the column was centrifuged at >10,000 rpm for 15 sec. The column was rinsed twice with 500 μL buffer RPE and centrifuged at >10,000 rpm for 15 sec and 2 min, respectively. The column was further centrifuged at 13,000 rpm for 1 min to remove any residual buffer RPE. 30 μL RNase-free water was added and the column was centrifuged at >10,000 rpm for 1 min. The eluted products was then re-applied to the column for further centrifugation at >10,000 rpm for 1 min. And finally, the total RNA was collected in a clean 1.5 mL centrifuge tube. The RNA yield and purity were determined spectrophotometrically with a BioPhotometer (Eppendorf AG, Hamburg, Germany) by measuring absorbance at 260nm and calculating the A_{260}/A_{280} ratio respectively. The extracted products were stored at -70 °C before further processing.

2.2.2.2 Converting *IDS* mRNA into cDNA by RT-PCR

cDNA was synthesized from the total RNA by RT-PCR. The reactions were carried out in the GeneAmp PCR System 9700 Cycler. 11 μL total RNA was mixed with 1 μL of 50 μM oligo dT and 1 μL of 10 mM dNTP Mix. The mixture was incubated at 65 °C in a heating block for 5 min and allowed to cool on ice for 1 min. A mixture consists of 4 μL first-strand buffer (5 ×), 1 μL of 0.1 M DTT and 2 μL SuperScript III RT was added to the cooled samples. The mixture was heated at 50 °C for 1 hour and the reaction was inactivated at 70 °C for 15 min. 1 μL RNase H was added and incubated with the mixture at 37 °C for 20 min. The products were stored at -20 °C before further processing.

2.2.2.3 Isolating *IDS* cDNA by PCR

IDS cDNA was first amplified by PCR using the first set of primer pairs in order to prevent non-specific products of pseudo-gene. Then, the PCR products were used as DNA templates for further PCR amplifying by the other sets of primer pairs. The first reactions were carried out in the MJ PTC-200 with the following parameters: one cycle of Taq polymerase activation at 94 °C for 12 min; 40 cycles of DNA denaturation at 94 °C for 30 sec, annealing at 60 °C for 45 sec, and extension at 72 °C for 2.5 min; one cycle of final extension at 72 °C for 10 min. The optimized annealing temperatures for the second set and the third set of primer pairs are 66 °C and 64 °C, respectively. And the optimized annealing temperatures for the forth set and fifth set of primer pairs are 65 °C. The reaction mixture of final volume 25 μL contained 1 μL cDNA template, 1 × PCR buffer, 50 nmol $MgCl_2$, 5 pmol dNTPs, 12.5 pmol of forward / reverse primers, and 0.625 U AmpliTaq Gold DNA polymerase. The PCR products were then analyzed by 2% agarose gel electrophoresis to ensure that the specific bands were in correct sizes.

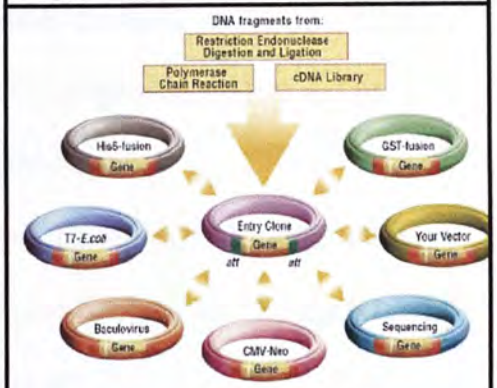
2.2.2.4 Isolating firefly *luciferase* gene by PCR

Luciferase cDNA was amplified by PCR using the control vector as DNA templates. The reactions were carried out in the MJ PTC-200 with the following parameters: one cycle of Taq polymerase activation at 94 °C for 12 min; 35 cycles of DNA denaturation at 94 °C for 30 sec, annealing at 60 °C for 45 sec, and extension at 72 °C for 2 min; one cycle of final extension at 72 °C for 10 min. The reaction mixture of final volume 25 μL contained 1 μL diluted control vector, 1 × PCR buffer, 50 nmol $MgCl_2$, 5 pmol dNTPs, 12.5 pmol of forward / reverse primers, and 0.625 U AmpliTaq Gold DNA polymerase. The PCR products were then analyzed by 2% agarose gel electrophoresis to ensure that the specific bands were in correct sizes.

2.3 Introducing *IDS* cDNA into Gateway Cloning System

One of the most common applications for PCR fragments is cloning into a plasmid vector for protein expression. Desired gene was cloned into pENTR/D-TOPO for entering into the Gateway system, which enabled enormous convenience in subsequent expression experiments using different system (Figure 2.2). All cloning procedures for *IDS* cDNA are identical to those for *luciferase* gene. Here only described the cloning procedures of *IDS* cDNA as example. There were two forms of *IDS* cDNA, the full length form and the N-terminally truncated form, cloned in this study.

Figure 2.2 Gateway technology



Adapted from Invitrogen, 2004

Gateway technology allows transferring DNA fragments into different expression vectors by lambda phage site-specific recombination [Landy A, 1989].

2.3.1 Materials

2.3.1.1 Directional cloning

pENTR Directional TOPO Cloning Kit (Invitrogen Corporation, Carlsbad, California, USA) provides pENTR TOPO reagents and TOP10 Chemically Competent *E.coli* set for directional cloning of PCR fragment. pENTR TOPO reagents include pENTR/D-TOPO vector, dNTP Mix, salt solution and sterile water. M13 Forward (-20) sequencing primer and M13 reverse sequencing primer are

included for verification of the insert by sequencing.

pENTR/D-TOPO vector is supplied in a concentration of 5-10 ng/μL. A detailed map is shown in Appendix 5.

pcDNA-DEST40 vector (Invitrogen Corporation, Carlsbad, California, USA) is supplied lyophilized in TE, pH 8.0. The vector is resuspended to 150 ng/ μL with RNase-free dH_2O. A detailed map is shown in Appendix 6.

pEXP1-DEST vector (Invitrogen Corporation, Carlsbad, California, USA) is supplied lyophilized in TE, pH 8.0. The vector is resuspended to 150 ng/ μL with RNase-free dH_2O. A detailed map is shown in Appendix 7.

QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, California, USA) consists of 10 × reaction buffer, QuikSolution reagent, dNTP mix, *Pfu*Ultra high fidelity DNA polymerase (2.5 U/ μL) and *Dpn*I restriction enzyme (10 U/ μL). It is stored at –20 °C. The kit claims to have an efficiency greater than 80%.

Mutagenic primers, containing the desired mutations of *IDS*, i.e. c.1016T>C and c.1016T>G, were designed by a web-based primer design software program at http://labtools.stratagene.com/QC. The primers were synthesized by QIAGEN (QIAGEN, Hilden, Germany) with HPLC purification. The primer sequences were described in Table 2.2.

LR Clonase Reaction mix (Invitrogen Corporation, Carlsbad, California, USA) consists of Gateway LR Clonase enzyme mix, LR reaction buffer and proteinase K solution. Gateway LR Clonase enzyme mix contains bacteriophage lambda recombination proteins Integrase and Excisionase, and the *E.coli*-encoded protein Integration Host Factor. It is stored at –70 °C.

TOP10 Chemically Competent E.coli set consists of SOC Medium and OneShot TOP10 Competent Cells. The TOP10 cells yield 1×10^9cfu/μgDNA in a 50 μL transformation reaction. The genotype of TOP10 is shown in Appendix 8.

2.3.1.2 LB medium/ agar with antibiotics preparation

Kanamycin sulfate (Invitrogen Corporation, California, USA) $C_{18}H_{36}N_4O_{11}\bullet H_2SO_4$, with MW 583, is supplied lyophilized. It is reconstituted to 10 mg/ mL with sterile dH_2O and stored at -20 °C. It is used as 50µg/ mL in LB medium or agar.

Ampicillin (Invitrogen Corporation, Carlsbad, California, USA) $C_{16}H_{18}N_3O_4SNa$, with MW 371.4, is supplied lyophilized. It is reconstituted to 10 mg/ mL with sterile dH_2O and stored at -20 °C. It is used as 100 µg/ mL in LB medium or agar.

LB Agar/ Lennox L agar (Invitrogen Corporation, Carlsbad, California, USA) is supplied in powder form and stored at RT. For preparation, 32 g was added to 1 L distilled water and autoclaved at 121 °C for 15 mins. The agar was cooled in a 55 °C water bath. Appropriate antibiotics were added. The agar was poured into plates and allowed to stand overnight. The plates were stored at 4 °C in dark.

LB Broth Base/ Lennox L Broth Base (Invitrogen Corporation, Carlsbad, California, USA) is supplied in powder form and stored at RT. For preparation, 20 g was added to 1 L distilled water and autoclaved at 121 °C for 15 mins. Appropriate antibiotics were added just prior to using the medium.

2.3.1.3 Plasmids purification from transformed cells

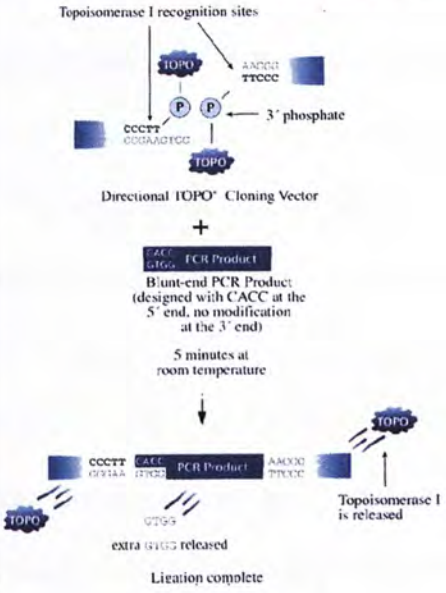
QIAprep Miniprep Kit (QIAGEN, Hilden, Germany) consists of buffer P1, buffer P2, buffer N3, buffer PB, buffer PE, buffer EB, RNase A, QIAprep spin columns and 2 mL collection tubes. RNase A is added to Buffer P1 before use. The ready-to-use buffer P1 is stored at 4 °C. All the other reagents are stored at RT. The DNA binding capacity of the membrane in a QIAprep spin column is 20 µg.

2.3.1.4 Validation of *IDS* inserted plasmids

Internal primers were designed by the web-based software program, OligoAnalyzer, and were synthesized by QIAGEN (QIAGEN, Hilden, Germany). Table 2.2 shows the sequences of the primers. Seven internal primers are used to screen the entire *IDS* cDNA: 1650 bp and 1551 bp. Sequencing using primer IDS-inter-R1 can cover the 5′-end of *IDS* cDNA while IDS-inter-F3 can cover the 3′-end. Similarly, internal primers were designed for sequencing the entire cDNA of *luciferase* (Table 2.3).

Sequencing reagents are the same as described in Chapter 2.1.4.1.

Figure 2.3 Gateway–directional TOPO cloning



Adapted from Invitrogen, 2004

TOPO cloning utilizes the enzyme topoisomerase I [Shuman S, 1994] replacing DNA ligase in traditional cloning. Topoisomerase I is isolated from vaccinia virus. It functions as both restriction enzyme and ligase. It specifically recognizes the sequence 5'-(C/T)CCTT-3' and covalently bind to the phosphate group of the 3'-thymidine. It cleaves one DNA strand, enabling the DNA to unwind. The enzyme then religates the ends of the cleaved strand and releases itself from the DNA.

2.3.2 Methods

2.3.2.1 TOPO cloning reaction

The sequence 5′-CACC-3′ was added to the *IDS* cDNA by specific primers either IDS-E-Fp or IDSt-E-Fp (Figure 2.3). 2 μL PCR products synthesized from the second and third sets of primer pairs were ligated with 0.5 μL pENTR/D-TOPO vector in 0.5 μL salt solution. The ligation mixtures were incubated at RT for 10 min. Then, the ligation products were

used for transformation immediately.

2.3.2.2 Transformation

The entire ligation products were added into 25 μL OneShot TOP10 Competent *E.coli* Cells. The mixture was incubated on ice for 30 min. This incubation allowed attachment of the plasmids onto the cell surface. The cells were heat shocked in a water bath at 42 °C for 30 seconds and immediately chilled on ice. This facilitated the ligation products entering into the competent cells. Then, the cells were recovered in 250 μL SOC medium by shaking at 37 °C, 200 rpm for 30 min. The transformed cells were then spread onto LB agar plates containing 50 μg/ mL kanamycin and incubated at 37 °C for 16 hours. A single colony was then picked and inoculated into 5 mL LB medium containing 50 μg/ mL kanamycin. The inoculum was shaken at 37 °C, 200 rpm for 16 hours. A total of 5 colonies were picked for subsequent experiments.

2.3.2.3 Small-scale plasmids preparation by QIAprep Miniprep Kit

The plasmid was extracted from the bacteria by using QIAprep Miniprep Kit according to the manufacturer's instruction (Figure 2.4). 5 mL of bacterial culture was pelleted by centrifugation at 3500 rpm for 10 min at 4 °C. The cell pellet was resuspended in 250 μL Buffer P1. And 250 μL Buffer P2 was added. The tube was inverted gently 4 times and incubated at RT for not more than 5 min to prevent release of genomic DNA. 350 μL Buffer N3 was then added and mixed as before. The mixture was centrifuged at 16100 × g for 10 min. The supernatant was applied to the QIAprep Spin Column and centrifuged at 16100 × g for 1 min. 500 μL Buffer PB was added and centrifuged at 16100 × g for 1 min. 750μL Buffer PE was added and centrifuged at 16100 × g for 1 min. The column was further centrifuged at

16100 × g for 1 min to remove any residual ethanol. The plasmids were eluted with 50 µL dH_2O. 1 min incubation at RT is prior to centrifugation at 16100 × g for 1 min. The DNA yield and quality were determined spectrophotometrically with a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, Delaware, USA) by measuring A_{260} and calculating the A_{260}/A_{280} ratio respectively.

2.3.2.4 Sequencing the plasmids

The cloning orientation and the entire sequence of *IDS* cDNA were verified by sequencing reaction. The purified plasmids were directly added as template. The reaction mixture of 10 µL final volume contains 2 µL purified plasmids, 1 pmol *IDS* internal primers, 2 µL BigDye Terminator v1.1 cycle sequencing mix and 1 µL of 5 × BigDye v1.1/3.1 Sequencing Buffer. The cycle reaction, purification of products and sequencing analysis were described as in Chaper 2.2.3.2. If there is no mutation found in the *IDS* cDNA and the insert is in desired direction in pENTR/D-TOPO vector, this entry clone will be stored at 4 °C for subsequence experiments.

Figure 2.4 QIAprep Miniprep Kit

Adapted from Qiagen, 2004

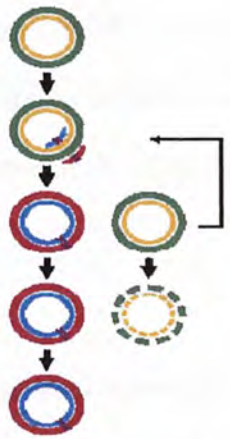
QIAprep Miniprep Kit utilizes a modified alkaline lysis method of bacteria cells [Birnboim HC and Doly J, 1979], in which Buffer P2 contains SDS, which solubilizes the phospholipids and protein components of the cell membrane and releases the cell contents. Buffer N3 was used to neutralize the lysate and its high salt concentration precipitates denatured proteins, chromosomal DNA, cellular debris, and SDS. The lysate is loaded into QIAprep columns with a silica-gel membrane which absorbs plasmid DNA in high-salt buffer and elutes in low-salt buffer.

2.3.2.5 QuikChange II XL site-directed mutagenesis

2.3.2.5.1 Synthesizing mutant strand with desired mutations

Mutagenesis was carried out in the entry clones to create the following amino acid changes: Leu339Pro and Leu399Arg (Figure 2.5). 10 ng dsDNA template was amplified in a reaction mix of 50 µL final volume containing 125 ng of the corresponding oligonucleotide primer and 125 ng of its anti-sense oligonucleotide primer, 1 µL dNTPs, 3 µL QuikSolution and 1x reaction buffer. An addition of 2.5 U *Pfu*Ultra HF DNA polymerase was then added. PCR parameters were as follow: denaturation at 95 °C for 1 min, 18 cycles of denaturation at 95 °C for 50 sec, primer annealing at 60 °C for 50 sec, extension at 68 °C for 5 min, and a final extension at 68 °C for 7 min.

Figure 2.5 QuikChange II XL Site-directed Mutagenesis Kit



Adapted from Stratagene, 2004

QuikChange II XL Site-Directed Mutagenesis Kit replaces, inserts, or deletes nucleotides from a double-stranded vector. Two synthetic oligonucleotide primers containing the desired mutation, each complementary to opposite strands of the vector, are extended to generate a mutated plasmid. The products are treated with *Dpn*I endonuclease (target sequence: 5'-Gm^{6}ATC-3'), which is specific for methylated and hemimethylated DNA [McClelland M and Nelson M, 1992]. The parental DNA template is dam methylated, thus it is digested. Only the mutant vector remains intact.

2.3.2.5.2 Digesting parental strand

The amplification products containing parental strand and mutated vector were digested by 10 U *Dpn*I restriction enzyme at 37 °C for 1 hour. 5 µL *Dpn*I-treated products were analyzed by 2 % agarose gel electrophoresis. A sharp band indicates successful mutant strand synthesis.

2.3.2.5.3 Transformation

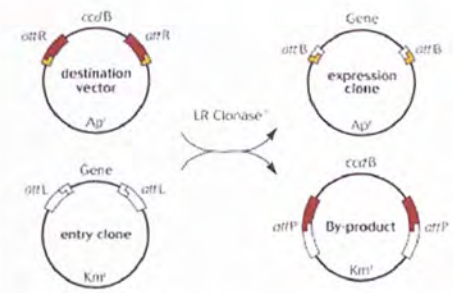
The intact mutant strand (2 μL) was transformed into OneShot TOP10 Competent Cells for propagation as described in Chapter 2.3.2.2. The plasmids were then purified by QIAprep Miniprep Kit as described in Chapter 2.3.2.3. Mutant *IDS* genes were obtained. The mutant plasmids were subjected to further sequencing to exclude the presence of undesired mutations in the *IDS* cDNA as described in Chapter 2.3.2.4.

2.3.2.6 Swapping *IDS* gene from entry clone to expression vectors

2.3.2.6.1 LR clonase reaction

Both the wild-type and mutant *IDS* cDNA were swapped from pENTR/D-TOPO into expression vectors by LR clonase reaction (Figure 2.6). 300 ng entry clones and 300 ng of the destination vector, pEXP1-DEST or pcDNA-DEST40, were recombined in a reaction mixture of 20 μL final volume containing 1 × LR reaction buffer and 4 μL Gateway LR clonase enzyme mix. The mixture was incubated in a water bath at 25 °C for 60 min. The reaction was terminated by the action of 2 μL proteinase K solution at 37 °C for 10 min.

Figure 2.6 LR clonase reaction



Adapted from Invitrogen, 2004

Gateway technology utilizes lambda phage site-specific recombination (attL × attR ⟷ attP × attB) to insert a gene into vector in replacement of the use of endonuclease and ligase in traditional cloning. Any DNA fragment flanked by a recombination site (attL) can be transferred into any vector that has a corresponding site (attR) by LR Clonase enzyme mix.

2.3.2.6.2 Transformation

The LR clonase reaction product (1 µL) was transformed into OneShot TOP10 Competent Cells as described in Chapter 2.3.2.2. For selecting the expression vectors from entry clones, the cells were cultured on LB agar plates / LB medium containing 100 µg/ mL ampicillin. The plasmids were then purified by QIAprep Miniprep Kit as described in Chapter 2.3.2.3. The expression clones were sequenced to exclude the presence of undesired mutations in the *IDS* cDNA as described in Chapter 2.3.2.4.

2.4 Introducing *IDS* cDNA into RTS pIVEX Wheat Germ vector

IDS cDNA was also cloned by traditional method. Restriction sites can be incorporated into PCR primers for standard sticky-ended cloning but this method introduces potential problems related to poor enzyme cleavage at DNA ends. Nevertheless, pIVEX_1.3_WG is highly recommended because it is specialized for in vitro expression in the Roche wheat germ-based cell-free system.

2.4.1 Materials

2.4.1.1 Restriction digestion

pENTR/D-TOPO vector containing IDS cDNA was cloned as described in Chapter 2.3.2. Only the entry vector carrying wild-type *IDS* cDNA was used.

pIVEX_1.3_WG vector (Roche Diagnostics GmbH, Basel, Switzerland) is supplied lyophilized in TE, pH 8.0. The vector is resuspended to 100 ng/µL with RNase-free dH_2O. A detailed map is shown in Appendix 9.

Nde I, 20,000 U/ mL (New England Biolabs, Ipswich, UK) is supplied with 10 × Buffer 4. The reagents are stored at -20 °C. The reagents are stored at -20 °C. The enzyme recognizes DNA sequence 5′-CATATG-′3 and forms sticky ends after cleavage.

Xho I, 20,000 U/ mL (New England Biolabs, Ipswich, UK) is supplied with 10 × Buffer 2 and 10 × Bovine Serum Albumin (BSA). The reagents are stored at -20 °C. The enzyme recognizes DNA sequence 5′-CTCGAG-′3 and forms sticky ends after cleavage.

2.4.1.2 Purification of digested products

QIAquick gel extraction kit (Qiagen, Hilden, Germany) comprises buffer QG, buffer PE, QIAamp spin columns, 2 mL collection tubes and 1.5 mL centrifuge tubes. All the reagents are stored at RT. The kit claims to yield DNA from 70 bp to 10 kb.

Isopropanol (Merck KGaA, Darmstadt, Germany) C_3H_8O, with MW 60.09, is supplied as colorless liquid with a pleasant odor. It is stored at RT.

2.4.1.3 Ligation of the *IDS* insert into pIVEX_1.3_WG

Ligate-IT rapid ligation kit (USB Corporation, Ohio, USA) contains Ligate-IT T4 DNA Ligase, 5 × Ligate-IT reaction buffer. The kit is stored at -20 °C.

Polyethylenglycol 1500 for synthesis/ PEG (Merck KGaA, Darmstadt, Germany) $HO(C_2H_4O)_nH$, with MW 1500, is stored at RT.

2.4.2 Methods

2.4.2.1 Restriction digestion to create sticky ends

IDS cDNA was amplified by the forth set of primer pairs as described in Chapter 2.2.2.3. Then, 7 µL *IDS* insert was digested by 10 U of *Nde I* and 10 U *Xho I* in a reaction mixture of 10 µL final volume containing 1 × Buffer 4 and 1 × BSA. 14.4 µL pIVEX_1.3_WG was digested by10 U of *Nde I* and 10 U of *Xho I* in a reaction mixture of 20 µL final volume containing 1 × Buffer 4 and 1 × BSA. The mixture was incubated at 37 °C for 2 hours. The digested products were analyzed by 2% agarose gel electrophoresis to ensure that the specific bands were in correct sizes.

2.4.2.2 Purifying the digested products

The digested products were purified from the agarose gel by QIAquick gel extraction kit. Both the specific bands of *IDS* and the vector were cut from appropriate positions on the agarose gel. 3 volume of buffer QG was added to 1 volume of gel slice (100 mg ~ 100μL). The gel was allowed to melt in 50 °C for 10 min. 1 volume of absolute isopropanol was added to the melted solution. Then the mixture was applied to the QIAprep Spin Column and centrifuged at 16100 × g for 1 min. 500 μL Buffer GC was added and centrifuged at 16100 × g for 1 min. 750μL Buffer PE was added and centrifuged at 16100 × g for 1 min. The column was further centrifuged at 16100 × g for 1 min to remove any residual ethanol. The DNA were eluted with 30 μL dH_2O. 1 min incubation at RT is prior to centrifugation at 16100 × g for 1 min. The DNA yield and quality were determined spectrophotometrically with a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, Delaware, USA) by measuring A_{260} and calculating the A_{260}/A_{280} ratio respectively.

2.4.2.3 Ligating the *IDS* insert into pIVEX_1.3_WG

The ligation reaction had 10 μL final volume containing 3 U of T4 DNA ligase, 1 × UBS buffer, 15 % PEG and insert to vector ratio = 3:1. The reaction was carried out in GeneAmp PCR System 9700 Cycler at 25 °C for 2 hours and 12 °C for 16 hours.

2.4.2.4 Transformation

The ligated product (5 μL) was transformed into OneShot TOP10 Competent Cells as described in Chapter 2.3.2.2. For selecting the expression vectors from entry clones, the cells were cultured on LB agar plates / LB medium containing 100 μg/

mL ampicillin. The plasmids were then purified by QIAprep Miniprep Kit as described in Chapter 2.3.2.3. Only the normal full length form of *IDS* was succefully cloned into the expression vector. The expression clones were sequenced to exclude the presence of undesired mutations in the *IDS* cDNA as described in Chapter 2.3.2.4.

2.5 Transient expression study of *IDS* constructs

2.5.1 Materials

Dulbecco's Modified Eagle Medium/ DMEM, powder (high glucose) (Invitrogen Corporation, Carlsbad, California, USA) is in standard formulation, containing 4500 mg/L D-glucose, and L-glutamine, but no sodium pyruvate or sodium bicarbonate. To prepare 1 × liquid medium, the powdered medium was dissolved in 1 L autoclaved dH_2O with 3.7 g $NaHCO_3$. The medium was adjusted to pH 7.1-7.2 with 1 N HCl or 1 N NaOH. Desired final working pH units will be achieved after filtration, usually pH will rise 0.1-0.3. In the biological safety cabinet, the medium was steriled by filtration with a 4.5 nm filter.

Fetal Bovine Serum, Certified (US)/ FBS (Invitrogen Corporation, Carlsbad, California, USA) is stored at -20 °C. It is tested to have endotoxin level lower than 10 EU/mL and hemoglobin level lower than 15 mg/dL. It is heat-inactivated at 55°C for 30 mins.

Phosphate Buffered Saline/ PBS (Sigma-Aldrich, St. Louis, Missouri, USA): as described in Chapter 2.2.1.1.

Hanks' Balanced Salt Solution/ HBSS (10 × liquid) (Invitrogen Corporation, Carlsbad, California, USA) is in standard formulation containing no calcium chloride, magnesium chloride, magnesium sulfate, or sodium bicarbonate. It is stored at 4 °C.

Trypsin-EDTA/ 0.5% Trypsin, EDTA•4Na (10 × liquid) (Invitrogen Corporation, Carlsbad, California, USA) contains 5.0 g /L trypsin (1:250), 2.0 g/L EDTA•4Na and 8.5 g/L NaCl. It was porcine parvovirus tested and mycoplasma tested. It was provided in 10 × solution and stored at -20 °C. To prepare 1 ×

Trypsin-EDTA, 10 mL Trypsin-EDTA was added to 90 mL HBSS.

Cell dissociation buffer, enzyme-free, PBS-based (Invitrogen Corporation, Carlsbad, California, USA) is stored at 4 °C.

Complete protease inhibitor cocktail tablets (Roche Diagnostics GmbH, Basel, Switzerland) is supplied as tablets and stored at 4 °C. To prepare working solution, one tablet is dissolved in 50 mL dH_2O.

Opti-MEMI Reduced Serum Medium (Invitrogen Corporation, Carlsbad, California, USA) contains HEPES buffer, 2400 mg/L sodium bicarbonate, hypoxanthine, thymidine, sodium pyruvate, L-glutamine, trace elements, growth factors, and 1.1mg/L phenol red. It is stored at 4 °C.

COS-7 cells (American Type Culture Collection, Manassas, Virginia, USA) are African Green Monkey, *Cercopithecus aethiops*, kidney cells in fibroblast morphology. It is supplied frozen on dry ice. Upon receipt, the cells were thawed quickly in a water bath at 37 °C. 10 mL DMEM medium with 10 % FBS was added dropwise. The cells were then centrifuged at 1000 rpm for 5 mins to remove the cryopreservative. The cells were resuspended in 9 mL DMEM medium with 10 % FBS and plated.

Lipofectamine 2000 (Invitrogen Corporation, Carlsbad, California, USA) is stored at 4 °C. The cationic lipid transfectant was mixed by gently flicking before use.

pcDNA-DEST40 vector containing IDS cDNA was cloned as described in Chapter 2.3.2. Transcription of *IDS* is under the control of a strong cytomegalovirus (CMV) promoter. The vector also consists of polyhistidine (6 × His) region at the 3′ end of *IDS* insert. However, the *IDS* cDNA was cloned with stop codon. Therefore, the IDS protein synthesized would not be His-tagged fusion protein.

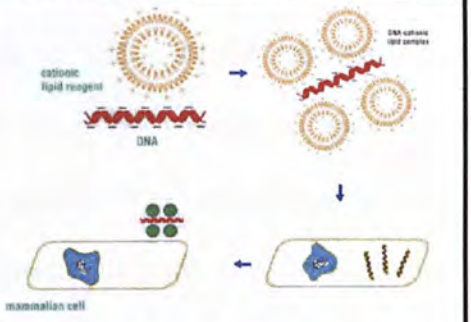
pcDNA-DEST40 vector containing luciferase cDNA was cloned as *IDS* cDNA

described in Chapter 2.3.2.

RNeasy Mini Kit (Qiagen, Hilden, Germany) is the same as described in Chapter 2.2.1.1.

RT-PCR and PCR reagents are the same as described in Chapter 2.2.1.2, Chapter 2.1.3.1.1 and Chapter 2.1.3.1.2.

Figure 2.7 Lipofectamine 2000



Adapted from Invitrogen, 2004

Cationic lipid reagent-mediated transfection [Felgner PL *et al.*, 1987] involves the formation of small unilamellar liposomes, which are positively charged and are attracted electrostatically to the negative phosphate backbone of DNA, forming DNA-cationic liposome complexes. The complexes are then attracted to the negatively charged surface of the cell membrane [Gershon H *et al.*, 1993] and enter the cells through endosomes and lysosomes into the cytoplasm [Coonrod A *et al.*, 1997].

2.5.2 Methods

2.5.2.1 Cell culturing

COS-7 cells were grown in DMEM medium, supplemented with 10% heat-inactivated FBS at 37 °C in 5% CO_2/air atmosphere.

2.5.2.2 Transfecting *IDS* constructs by lipofection procedures

All transfections were performed in triplicate. One day before transfection, 0.48×10^6 COS-7 cells were seeded onto one of the wells of the 6-well culturing plate for each samples. Introduction of the vector pcDNA-DEST40 containing *IDS* cDNA into the COS-7 cells was performed by using Lipofectamine 2000 (Figure 2.7). The vector pcDNA-DEST40 containing *luciferase* cDNA was co-transfected with either the wild-type *IDS* or the mutant *IDS* as control for transfectioin efficiency and expression efficiency. 0.96 μg of plasmid containing *IDS* cDNA and

0.64 μg of plasmid containing *luciferase* cDNA were diluted in 100 μL Opti-MEMI Reduced Serum Medium. 4 μL Lipofectamine 2000 was diluted with 100 μL Opti-MEMI Reduced Serum Medium and the mixture was incubated at RT for 5 min. Diluted Lipofectamine 2000 and the plasmids were combined together and incubated at RT for 20 min. In the meantime, the overnight cultures of COS-7 cells were rinsed with PBS twice to remove any FBS which may interference with transfection. Then the cells were supplied with 400 μL plain DMEM medium and 200 μL Opti-MEMI Reduced Serum Medium. The DNA-cationic liposome complexes were added to the cells dropwise. The mixture was incubated at 37 °C with 5% CO_2 for 6 hours. Then the medium was replaced by 2 mL DMEM medium supplemented with 10 % FBS. The transfected cells were incubated at 37 °C with 5% CO_2 for further 42 hours.

2.5.2.3 Harvesting COS-7 cells

The transfected COS-7 cells were harvested after incubation for 48 hours in total. The culturing medium was collected and stored at -20 °C. 0.5 mL cell dissociation buffer was added to the transfected cells and incubated 37 °C with 5% CO_2 for 15 min. 1 mL PBS was added to rinse and detach the cells from the plate. The cells were collected and centrifuged at 2,500 rpm at 4 °C for 5 min. Then the cell pellets were rinse with 1 mL PBS and subjected to further centrifugation at 2,500 rpm at 4 °C for 5 min. The cell pellets were resuspended with 50 μL dH_2O with addition of protease inhibitor. The COS-7 cells were then stored at -70 °C for 1 hour before sonication. Since thermal denaturation of IDS revealed that a significant loss of enzyme activity above 50 °C, the whole process of sonication was performed on ice to prevent over-heating [Parkinson-Lawrence E *et al.*, 2005]. There were four rounds of sonication with each of them lasted for 5 sec. Each rounds were performed

every 30 sec. The lysate of transfected COS-7 cells were centrifuged at 13,000 × g at 4 °C for 5 min. The supernatants were stored at -20 °C for subsequence experiments.

2.5.2.4 Total RNA extraction from transfected COS-7 cells

IDS tranfected COS-7 cells were prepared as described in Chapter 2.5.2.2. The cells were harvested 4 hours after transfection. 0.2 mL PBS with 10 % Trypsin-EDTA was added to the cells and incubated at 37 °C for 3 min. 0.4 mL PBS was added to rinse and detach all the cells. The cells were centrifuged at 300 × g for 5 min at 4 °C. Total RNA was isolated from the cell pellet as described in Chapter 2.2.2.1 with some modifications. The cell pellet was loosened from the bottom by flicking prior to resuspension with 350 µL buffer RLT. The transfected COS-7 cells were homogenized through the QIAShredder Spin Column by centrifugation at 13,000 rpm for 2 min. 350 µL of 70 % ethanol was mixed with the lysate. And the mixture was applied to the RNeasy mini column. The cell debris was filtered by centrifugation at >10,000 rpm for 15 sec and discarded. 700 µL buffer RW1 was added, and the column was centrifuged at >10,000 rpm for 15 sec. The column was rinsed twice with 500 µL buffer RPE and centrifuged at >10,000 rpm for 15 sec and 1 min, respectively. The column was further centrifuged at 13,000 rpm for 1 min to remove any residual buffer RPE. 30 µL RNase-free water was added and the column was centrifuged at >10,000 rpm for 1 min. The eluted products was then re-applied to the column for further centrifugation at >10,000 rpm for 1 min. And finally, the total RNA was collected in a clean 1.5 mL centrifuge tube. The RNA yield and purity were determined spectrophotometrically with NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, Delaware, USA) by measuring absorbance at 260nm and calculating the A_{260}/A_{280} ratio respectively. The extracted products were stored at -70 °C before further processing.

2.5.2.5 RT-PCR showing *IDS* mRNA stability

RT-PCR was performed as described in Chapter 2.2.2.2. *Luciferase* cDNA was also amplified for normalization by PCR as described in Chapter 2.2.2.3. The flanking primer pairs used were Luci-E-Fp and Luci-InterR2. The optimized annealing temperature was 60 °C. The appropriate cycling number was found to be 24 cycles. Then fragments of *IDS* cDNA were amplified by PCR as described in Chapter 2.2.2.3 but using 24 cycles only. There were two sets of primer pairs, i.e. **1)** IDS-E-Fp and IDS-InterR3, **2)** IDS-InterF2 and IDS-InterR3, used for amplifying different regions of *IDS* cDNA. The optimized annealing temperature for both primer pairs was 64 °C. The PCR products were then analyzed by 2% agarose gel electrophoresis to ensure that the specific bands were in correct sizes. The density of each DNA bands were compared by using BIO RAD Quantity One 4.2.1 Model GS-700 Imaging Densitometer (Bio-Rad Laboratories, Hercules, California, USA).

2.5.2.6 Endocytosis of expressed IDS products into COS-7 cells

One day before endocytosis, 0.48×10^6 COS-7 cells were seeded onto one of the wells of the 6-well culturing plate for each samples. The culturing medium was replaced by 1.5 mL fresh DMEM medium supplemented with 10 % FBS. 500 μL expression products of IDS were added dropwise. For the IDS synthesized by transfected COS-7 cells, 500 μL cell-free medium was used as sample. The medium of untransfected cells was used as negative control. For the IDS synthesized by cell-free expression systems, 15 μL expression products were diluted with 485 μL DMEM medium before added to the cells. The cell-free expression reagents without DNA template were used as negative control. COS-7 cells immersed with IDS products were incubated at 37 °C with 5% CO_2. Then the cells were harvested immediately or after 48-hour incubation as described in Chapter 2.5.2.3.

2.6 Synthesizing IDS by cell-free *in vitro* expression systems

2.6.1 Materials

2.6.1.1 DNA templates for expression

pEXP1-DEST vector containing IDS cDNA (Invitrogen Corporation, Carlsbad, California, USA) was cloned as described in Chapter 2.3.2. Transcription of *IDS* is under the control of T7 promoter. The expression vector also consists of polyhistidine (6 × His) region at the 5′ end of *IDS* insert. As the cell-free systems do not support signal peptide cleavage, the IDS protein synthesized would be His-tagged fusion protein at the N-terminal end.

pIVEX_1.3_WG vector containing IDS cDNA (Roche Diagnostics GmbH, Basel, Switzerland) was cloned as described in Chapter 2.4.2. Transcription of *IDS* is under the control of T7 promoter. The expression vector also consists of polyhistidine (6 × His) region at the 5′ end of *IDS* insert. As the cell-free systems do not support signal peptide cleavage, the IDS protein synthesized would be His-tagged fusion protein at the N-terminal end.

pEXP1-DEST vector containing luciferase cDNA (Invitrogen Corporation, Carlsbad, California, USA) was cloned as described in Chapter 2.3.2. It served as positive control and has the advantage to be detected quantitatively by activity assay.

Control vector GFP (Roche Diagnostics GmbH, Basel, Switzerland) is supplied as lyophilized in TE, pH 8.0. The vector is resuspended to 1 μg/ μL with RNase-free dH_2O. The coding sequence of green fluorescent protein (GFP) is inserted in the pIVEX expression vector under the control of T7 promoter. It serves as positive control in bacterial-based expression systems and has the advantage to be detected easily by visualizing under fluorescence light.

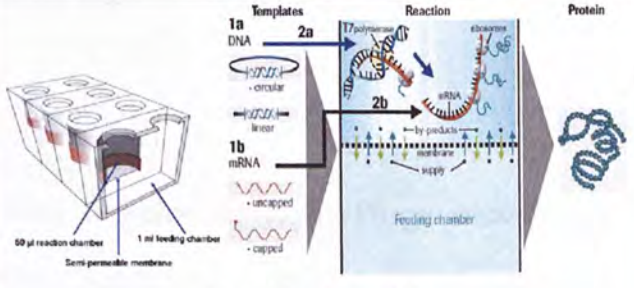
GUS Control vector (Roche Diagnostics GmbH, Basel, Switzerland) is supplied as lyophilized in TE, pH 8.0. The vector is resuspended to 1 µg/ µL with RNase-free dH_2O. The coding sequence of glucuronidase (GUS) is inserted in the pIVEX expression vector under the control of T7 promoter. It served as positive control in wheat germ-based expression system.

2.6.1.2 Commercial cell-free expression kits (Table 2.4)

ExpressWay Plus Expression System (Invitrogen Corporation, Carlsbad, California, USA) consists of IVPS plus *E. coli* extract, 2.5 × IVPS Plus *E. coli* reaction buffer, T7 enzyme mix, 75 mM methionine, DNase/RNase-free distilled H_2O, RNase A (1mg/ mL in DNase/RNase-free H_2O) and 2 mL screw-cap tubes. All the reagents are stored at -80 °C. The kit claims to have a production of up to 50 µg protein in 50 µL reaction.

RTS 100 E.coli HY Kit (Roche Diagnostics GmbH, Basel, Switzerland) consists of *E.coli* lysate, reaction mix, amino acids, methionine and reconstitution buffer. Most reagents are supplied as lyophilized powder and need to be reconstituted by the reconstitution buffer before use. All the reagents are stored at -20 °C. The kit claims to have a production of up to 20 µg protein in 50 µL reaction.

Figure 2.8 Microplate specialized for CECF expression



Adapted from Roche, 2004

Continuous exchange cell-free (CECF) expression: the presence of the semi-permeable membrane allows separation of the lower feeding chamber and the upper reaction chamber which facilitates supply of substrates and removal of inhibitors continuously.

RTS 100 Wheat Germ CECF Kit (Roche Diagnostics GmbH, Basel, Switzerland) consists of wheat germ lysate, reaction mix, feeding mix, amino acids, methionine, reconstitution buffer and microplate with adhesive film (Figure 2.8). Most reagents are supplied as lyophilized powder and need to be reconstituted by the reconstitution buffer before use. All reagents are stored at -20 °C except the reconstituted lysate which is stored at -80 °C. The microplate is keep at RT. The kit claims to have a production of over 20 µg protein in 50 µL reaction.

TNT Coupled Wheat Germ Extract Systems (Promega Corporation, Madison, USA) consists of TNT wheat germ extract, TNT reaction buffer, TNT T7 RNA polymerase and three types of 1mM amino acid mixture (minus methionine, leucine or cysteine). All the reagents are stored at -80 °C.

°C. The kit claims to have a production of up to 500 ng protein in 50µL reaction.

TNT Coupled Reticulocyte Lysate Systems (Promega Corporation, Madison, USA) consists of TNT rabbit reticulocyte lysate, TNT reaction buffer, TNT T7 RNA polymerase and three types of 1mM amino acid mixture (minus methionine or leucine or cysteine). All the reagents are stored at -80 °C. The kit claims to have a production of up to 500 ng protein in 50µL reaction.

2.6.1.3 Supplements

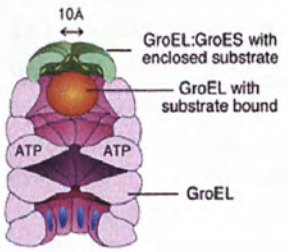
Rapid Translation System RTS GroE Supplement (Roche Diagnostics GmbH, Basel, Switzerland) is stored at -20 °C. It consists of GroEL and GroES which are called molecular chaperones. The chaperones accomplish protein folding by directly interaction with the polypeptide. However, > 30 % synthesized polypeptides cannot fold correctly to form functional proteins even having the chaperone machinery.

Canine Pancreatic Microsomal Membranes (Promega Corporation, Madison, USA) is stored at -80 °C. It consists of microsomal vesicles which are isolated free

from contaminating membrane fractions and stripped of endogenous membrane-bound ribosomes and mRNA. Addition of the microsomes helps post-translational processing of proteins such as signal peptide cleavage and glycosylation but diminishes 10 - 50 % translation efficiency.

RNasin Ribonuclease inhibitor (Promega Corporation, Madison, USA) is supplied as 40U/ µL and is stored at -20 °C.

Figure 2.9 Rapid translation system RTS GroE supplement



Adapted from Roche, 2004

Rapid Translation System RTS GroE Supplement contains GroEL and GroES. GroEL is arranged in 2 heptameric rings which form a cylindric complex. GroES forms dome-shaped hepatmeric rings which cap the GroEL cylinder on one or both ends. The size of the cylinder allows proteins with a molecular mass up to 60kDa to be substrates for the GroE system.

Table 2.4 Differences between the commercial cell-free expression systems

	Expressway Plus system	RTS 100 E.coli HY	RTS 100 WG.CECF	TnT coupled WG	TnT coupled Recticulocyte
Supplier	Invitrogen	Roche	Roche	Promega	Promega
Translation machinery	prokaryotic	prokaryotic	eukaryotic	eukaryotic	eukaryotic
Cell extracts	*E. coli*	*E. coli*	wheat germ	wheat germ	Rabbit reticulocyte
RNA polymerase	T7	T7	T7	T7	T7
Glycosylation/ signal peptides removal	N/A	N/A	N/A	N/A	N/A
Supplements	GroE	-	-	-	Microsomal vesicles
Functions of supplements	Refolding the protein	-	-	-	Assist the glycosylation / signal peptides cleavage
Reaction conditions	30 °C 4 hours	30 °C 6 hours	24 °C 24 hours	30 °C 1.5 hours	30 °C 1.5 hours
Max. protein products	50 µg per 50 µL	20 µg per 50 µL	> 20 µg per 50 µL	500 ng per 50 µL	500 ng per 50 µL

Table 2.4 Summary of the similarities and the differences between the five cell-free *in vitro* expression systems. From our experience, the addition of GroE supplements inhibited overall expression in the Roche *E. coli*-based system; and the addition of microsomal vesicles inhibited overall expression in the Roche wheat germ-based system. N/A means not available. "-" means not applicable.

2.6.2 Methods

2.6.2.1 Cell-free expression by ExpressWay plus expression system

The cell-free system is coupled with transcription and translation in a single reaction tube (Figure 2.10). 1 μg plasmid DNA was expressed in a reaction mixture of 56 μL final volume containing 20 μL IVPS plus *E.coli* extract, 20 μL 2.5× IVPS plus *E.coli* reaction buffer, 1 μL T7 enzyme mix, 1 μL 75mM methionine, 1 μL RNasin Ribonuclease inhibitor and 6 μL RTS GroE Supplement. The mixture was shaken in RTS ProteoMaster Instrument (Roche Diagnostics GmbH, Basel, Switzerland) at 300 rpm, 30 °C for 4 hours. 5 μL RNase A was added and the mixture was shaken for further 15 min at 37 °C. GFP or luciferase was expressed simultaneously as positive controls. Expressed GFP requires 24 hours incubation at 4 °C for maturation prior to fluorescent detection.

2.6.2.2 Cell-free expression by RTS 100 E.coli HY Kit

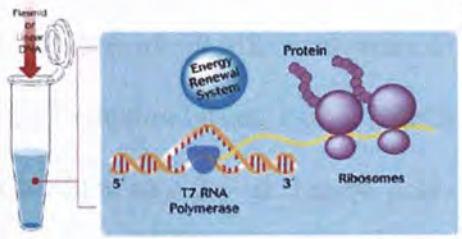
The cell-free system applied similar expression machinery as described in Figure 2.10. It was found that addition of RTS GroE Supplement inhibited the overall protein expression in this system. 1 μg plasmid DNA was expressed in a reaction mixture of 50 μL final volume containing 12 μL *E.coli* extract, 10 μL reaction buffer, 12 μL amino acids, 1 μL methionine, 1 μL RNasin Ribonuclease inhibitor and 5 μL reconstitution buffer. The mixture was incubated at 30 °C for 6 hours. GFP or luciferase was expressed simultaneously as positive controls.

2.6.2.3 Cell-free expression by RTS 100 Wheat Germ CECF Kit

Protein expression is in a two-chamber reaction format. Feeding solution and reaction solution were prepared independently. 1 mL feeding solution containing

900 µL feeding mix, 80 µL amino acids, 20 µL methionine and 0.5 µL RNasin Ribonuclease inhibitor was filled into the feeding chamber. 2 µg plasmid DNA was expressed in a reaction mixture of 50 µL final volume containing 15 µL wheat germ lysate, 15 µL reaction mix, 4 µL amino acids, 1 µL methionine, 0.5 µL RNasin Ribonuclease inhibitor and 5 µL reconstitution buffer. The mixture was shaken in RTS ProteoMaster Instrument (Roche Diagnostics GmbH, Basel, Switzerland) at 900 rpm, 24 °C for 24 hours. GUS was expressed simultaneously as positive control.

Figure 2.10 ExpressWay plus expression system



Adapted from Invitrogen, 2004

ExpressWay Plus Expression System utilizes extract from an S30 *E.coli* strain for expression of target DNA to protein [Zubay G, 1973]. The cell extract provides cellular machinery which is coupled with *in vitro* transcription and translation. The IVPS reaction buffer contains an ATP regenerating system for energy and all the required amino acids except methionine [Lesley SA, Brow MA and Burgess RR, 1991]. The bacterial systems have advantage of fast protein expression. The 5' end of the transcribed mRNA is bound to ribosomes and undergoes translation, while the 3' end of the template is still being transcribed.

2.6.2.4 Cell-free expression by TNT Coupled Wheat Germ Extract Systems

Similar to *E.coli*-based system, the wheat germ extract contains cellular machinery for *in vitro* transcription and translation in a single reaction tube. 1 µg plasmid DNA was expressed in a reaction mixture of 50 µL final volume containing 25 µL TNT wheat germ extract, 2 µL TNT reaction buffer, 1 µL T7 RNA polymerase, 0.5 µL amino acids mixture (minus methionine), 0.5 µL amino acids mixture (minus leucine) and 1 µL RNasin Ribonuclease inhibitor. The mixture was incubated at 30 °C for 90 min. Luciferase was expressed simultaneously as positive control.

2.6.2.5 Cell-free expression by TnT Coupled Reticulocyte Lysate Systems

The use of reticulocyte lysate was increased to 55 % of total reaction volume. 1 μg plasmid DNA was expressed in a reaction mixture of 50 μL final volume containing 27.5 μL TnT wheat germ extract, 2 μL TnT reaction buffer, 1 μL T7 RNA polymerase, 0.5 μL amino acids mixture (minus methionine), 0.5 μL amino acids mixture (minus leucine) and 1 μL RNasin Ribonuclease inhibitor. The mixture was incubated at 30 °C for 90 min. Luciferase was expressed simultaneously as positive control.

2.7 Investigations of IDS protein expression

The presence of expressed IDS was examined by simple protein staining after SDS-PAGE. For those IDS synthesized from the cell-free systems, the results were further confirmed by staining His-tagged proteins in the expression products. To determine the catalytic activity of the expressed IDS, a fluorometric enzyme assay was used.

2.7.1 Materials

2.7.1.1 Isolation of Histidine-tagged proteins

Dynabeads TALON solution (Dynal Biotech ASA, Oslo, Norway) is stored at 4 °C.

TALON binding buffer and washing buffer (Dynal Biotech ASA, Oslo, Norway) is stored at 4 °C.

TALON elution buffer (Dynal Biotech ASA, Oslo, Norway) is stored at 4 °C.

Magnetic stand (Dynal Biotech ASA, Oslo, Norway) is suitable for 1.5 mL centrifugation tubes.

2.7.1.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis/ SDS-PAGE

Sodium Dodecyl Sulfate/ SDS (Sigma-Aldrich Co., St. Louis, Missouri, USA) is supplied in powder form. 10% SDS is prepared by dissolving 10 g SDS with 100 mL dH_2O. It is stored at RT.

30% Acrylamide/Bis solution (Bio-rad Laboratories, Hercules, California, USA) is stored at 4 °C in dark.

Ammonium persulphate/ APS (Sigma-Aldrich Co., St. Louis, Missouri, USA) is supplied in powder form. 10 % APS solution is prepared by dissolving 0.5 g APS with 5 mL dH_2O. Dissolved APS is stored at -20 °C.

N,N,N',N'-tetramethylenediamine/ TEMED (Sigma-Aldrich Co., St. Louis, Missouri, USA) is stored at 4 °C in dark.

Tris Base (Sigma-Aldrich Co., St. Louis, Missouri, USA) is supplied in powder form. Tris-HCl buffer was prepared by dissolving Tris-base with dH_2O to a desired concentration and adjusting the pH with HCl. Tris-HCl buffer was stored at 4 °C.

10 % running gel was prepared by mixing 4 mL water, 3.33 mL 30 % acrylamide/Bis solution, 2.5 mL of 1.5 M Tris-HCl (pH 8.8), 50 µL 10 % SDS, 150 µL 10 % APS and 2.5 µL TEMED.

4.5 % stacking gel was prepared by mixing 3.4 mL water, 0.9 mL 30 % acrylamide/Bis solution solution, 1.5 mL 0.5 M Tris-HCl (pH 6.8), 60 µL 10 % SDS, 80 µL 10 % APS and 6 µL TEMED.

10× Tris-Glycine-SDS Electrophoresis Buffer is prepared by dissolving 0.25 M Tris-HCl (pH 7.5), 1.92 M glycine and 1 % SDS. It is stored at 4 °C. It is diluted to 1 × working solution with dH_2O.

Sample buffer is prepared by mixing 25 mL Tris-HCl (pH 6.8), 15.83 mL glycerol, 10 g SDS, 0.005 % bromophenol blue, 5 mL 20 mM EDTA and adjusting the pH to 7.5. 9 volumes of sample buffer are diluted to working solution with 1 volume of beta-mercaptoethanol.

Kaleidoscope Prestained Standards (Bio-Rad Laboratories, Hercules, California, USA) contains approximately 1.6 mg total proteins and acts at protein size marker. It is stored at -20 °C. Before use, it is heated at 40 °C for 10 min to separate any aggregated proteins. A detailed description is shown in Appendix 4.2.

Acetic acid, glacial 100 % GR (Merck KGaA, Darmstadt, Germany)

CH_3COOH, has a MW 60.05. It is corrosive and should be stored in fume cupboard.

Methanol, >99.8 % GR (Merck KGaA, Darmstadt, Germany) CH_3OH, has a MW 32.04. It is highly flammable and toxic and should be stored in fume cupboard.

Fixing solution/ destaining solution is prepared by mixing acetic acid, methanol and dH_2O in ratio of 1:3:10. It is stored at RT

Coomassie Brilliant Blue R-250 staining solution, 1L (Bio-Rad Laboratories, Hercules, California, USA) is stored at 4 °C.

Staining solution is prepared by adding 0.05 % Coomassie Blue in fixing solution. It is stored at RT.

InVision His-tag In-gel Staining Kit (Invitrogen Corporation, Carlsbad, California, USA) consists of InVision His-tag In-gel Stain. BenchMark His-tagged Protein Standard is included as the marker and positive control of the kit. The kit claims to detect 0.5 pmol of a 6 x His-tagged fusion protein. The stain is stored at RT but BenchMark His-tagged Protein Standard is stored at -20 °C. A detailed description of BenchMark His-tagged Protein Standard is shown in Appendix 4.3.

Potassium phosphate (Dibasic) (Sigma-Aldrich, St. Louis, Missouri, USA) K_2HPO_4 has a MW 174.2, anhydrous (98 %). It is in powder form and stored at RT.

Potassium phosphate (Monobasic) (Sigma-Aldrich, St. Louis, Missouri, USA) KH_2PO_4 has a MW 136, anhydrous (99+ %). It is in powder form and stored at RT.

2.7.1.3 Fluorometric activity assay for IDS

Sodium acetate, 99 % (Sigma-Aldrich, St. Louis, Missouri, USA) $C_2H_3O_2Na$, with MW 82.03, is supplied in powder form. To prepare 0.1 M solution, 8.2 g of powder is dissolved in 1 L dH_2O. It is stored at RT.

Acetic acid, glacial 100 % GR (Merck KGaA, Darmstadt, Germany) CH_3COOH, with MW 60.05, is stored in fume cupboard at RT. To prepare 0.1 M

solution, 5.775 mL of 100% acetic acid is diluted to 1 L with dH_2O inside the fume cupboard.

LeadII acetate trihydrate GR ACS ISO 99.5+ % (International Laboratory Limited, Nevada, USA) $C_4H_6O_4Pb.3H_2O$, with MW 379.33, is supplied in powder form. It is stored at RT.

Substrate buffer, pH 5.0 is prepared by adding 3.79 mg lead acetate into the 100 mL solution mixture containing 35.2 mL of 0.1 M sodium acetate and 14.8 mL of 0.1 M acetic acid. It is stored at RT.

4-methylumbelliferyl-α-L-iduronide-2-sulphate.Na2/ 4-Mu-αIdu-2S (Dept. Clinical Genetics, Netherlands) is supplied as lyophilized powder. To prepare 1.25 mM substrate solution, 0.6 mg powder is dissolved in 1 mL substrate buffer. The substrate solution is stored at -80 °C.

7-Hydroxy-4-methylcoumarin/ 4-Methylumbelliferone/ 4-Mu (International Laboratory Limited, Nevada, USA) $C_{10}H_8O_3$, with FW 176.17, is supplied as powder form and stored at 4 °C. It is dissolved in substrate buffer to produce a series concentration of standard 4-Mu from 0.2 µM to 1 M.

Citric acid, monohydrate (Sigma-Aldrich, St. Louis, Missouri, USA) $C_6H_8O_7.H_2O$, with FW 210.1 is supplied in powder form. To prepare 0.2 M solution, 38.42 g powder is dissolved in 1 L dH_2O. It is stored at RT.

Sodium phosphate, dibasic anhydrous, reagent grade (Sigma-Aldrich, St. Louis, Missouri, USA) Na_2HPO_4, with FW 142 is supplied in powder form. To prepare 0.4 M solution, 56.78 g powder is dissolved in 1 L dH_2O. It is stored at RT.

Sodium azide/ Na-azide (Sigma-Aldrich, St. Louis, Missouri, USA) NaN_3, with MW 65.01 is supplied in powder form and stored in RT. It is very toxic.

Pi/Ci-buffer, pH 4.5 (double concentrated McIlvains phosphate/ citric-acid buffer) is prepared by adding 0.02 % Na-azide into the 100 mL solution mixture

containing 26.7 mL of 0.2 M citric acid and 23.3 mL of 0.4 M Na_2HPO_4. It is stored at RT.

Sodium hydrogen bicarbonate (May & Bayer Ltd, Dagenham, UK) $NaHCO_3$, with MW 84.01, is supplied in powder form. To prepare 0.5 M solution, 53 g of powder is dissolved in 1 L dH_2O. It is stored at RT.

Sodium carbonate (Sigma-Aldrich, St. Louis, Missouri, USA) Na_2CO_3, with MW 106, is supplied in powder form. To prepare 0.5 M solution, 42 g of powder is dissolved in 1 L dH_2O. It is stored at RT.

Triton X-100 (BDH Chemicals Ltd, Poole, UK) $C_{34}H_{62}O_{11}$ with MW 647, is stored at RT.

Stop buffer, pH 10.7 is prepared by adding 0.025 % Triton X-100 into the 100 mL solution mixture containing 87.9 mL of 0.5 M $NaHCO_3$ and 12.1 mL of 0.5 M Na_2CO_3. It is stored at RT.

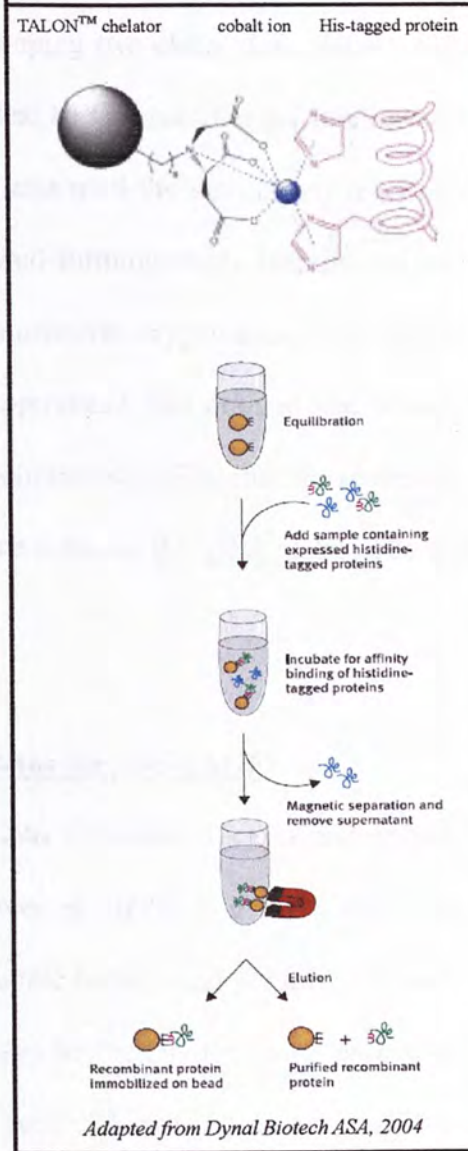
Bovine Serum Albumin/ BSA (Sigma-Aldrich, St. Louis, Missouri, USA) with MW 66,000 is stored at -20 °C. It is pH and heat denatured as follows: 5 g BSA powder is dissolved slowly in 40 mL dH_2O to form a 10 % solution. Then the pH value of dissolved BSA was adjusted to pH 10.0 with 6 M sodium hydroxide. It is heated at 50 °C for 4 hours. The pH value of cooled solution was then adjusted to pH 7.0 with 1 M HCl. Final volume of the neutral BSA was adjusted to 50 mL with dH_2O. Precipitates were removed by centrifugation at 3000 rpm for 10 min. 0.02 % of Na-azide was added to the supernatant. Inactivated BSA solution is diluted 50-fold to 0.2 % and stored at -20 °C.

Lysosomal enzymes/ LEBT (Dept. Clinical Genetics, Netherlands) is supplied as lyophilized powder. The enzymes mix is purified form bovine testis. To prepare working solution of LEBT, one vial of powder is reconstituted with 2.2 mL dH_2O and stored in aliquots at -80 °C.

2.7.1.4 Luciferase activity assay

Luciferase assay reagent (Promega Corporation, Madison, USA) consists of Luciferase Assay Substrate dissolved in Luciferase Assay Buffer. It is stored in aliquots at -80 °C.

Figure 2.11 Dynabeads TALON



Adapted from Dynal Biotech ASA, 2004

Purification is based on the metal coordination of the histidine-tag to the cobalt ion held in the TALON chelator.

2.7.2 Methods

2.7.2.1 Isolating His-tagged IDS from cell-free expression products

It is magnetic bead-based protein isolation. 50 µL Dynabeads TALON solution (2 mg) was equilibrated in 700 µL binding and washing buffer. The beads were pelleted by using a magnet and the supernatant was discarded. 700 µL binding and washing buffer containing 20 µL cell-free expression products were added. The mixture was incubated on a roller for 10 min at RT. By using the magnet, the beads were rinsed with 700 µL binding and washing buffer for 4 times. Then, 100 µL elution buffer was added and incubated for 5 min at RT. The supernatant was collected.

2.7.2.2 Protein staining of expression products

2.7.2.2.1 Preparation of protein separating gel

Bio-Rad Mini-Protein II was set up by clamping two clean glass plates which were held apart by plastic spacers on the two sides. 10 % separating gel solution was pipetted into the space between the two glass plates until the liquid level reached a height of 1 cm below the lowest portion of the well-forming comb. Isopropanol was added to cover the meniscus in order to keep atmospheric oxygen away from the gel solution. Upon setting of the separating gel, isopropanol was drained and washed away with deionized water. 4.5 % stacking gel solution was filled into the remaining space. Well-forming comb was put into the space between the glass plates. The gel was allowed to stand at RT until set.

2.7.2.2.2 Preparation of proteins for SDS-PAGE

The expression products were separated into insoluble fraction and soluble fraction of protein by centrifuging 5 μL samples at 10000 × g for 8 min. The samples were resuspended in 20 μL 1 × SDS sample buffer. And the mixtures were heated at 98 °C for 3 min. The protein products synthesized by the two *E.coli*-based cell-free expression systems required additional purification step before adding SDS sample buffer. 20 μL of ice-cool acetone was added to the supernatant and protein pellet. The mixtures were incubated on ice for 5 min for protein precipitation. Then the proteins were air-dried after centrifugation at 12000 rpm for 5 min.

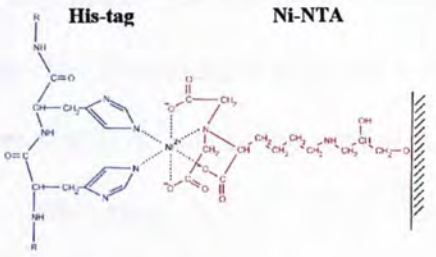
2.7.2.2.3 SDS-PAGE analysis

The proteins were separated by SDS-PAGE. 6 μL insoluble samples and 3 μL soluble samples were loaded per lane. 3 μL Kaleidoscope Prestained Standards was loaded as marker. After electrophoresis, the gel was immersed in 15 mL staining

solution and shaken at 100 rpm for 1 hour. Then the gel was destained by immersing in 15 mL staining solution and shaking at 100 rpm for 1 hour.

Fixed proteins can be stained by InVision his-tag In-gel stain (Figure 2.11). 5 μL BenchMark His-tagged Protein Standard was loaded as marker and positive control. After electrophoresis, the gel was immersed in 15 mL fixing solution for 1 hour. Fixed proteins were rinsed with 100 mL dH_2O twice for 10 minutes to remove the fixative. The gel was then immersed in 15 mL InVision His-tag In-gel Stain for 2 hours. The gel was rinsed with 15 mL 20 mM Phosphate Buffer (pH 7.8) for 10 minutes. Stained proteins were visualized using a 302 nm UV transilluminator equipped with an ethidium bromide filter and the fluorescent image was captured with a CCD camera. The exposure time was 30 seconds.

Figure 2.12 InVision His-tag In-gel Stain



Adapted from Qiagen, 2003

The InVision His-tag In-gel Stain is a stain capable of visualizing His-tagged fusion proteins in a polyacrylamide gel after electrophoresis directly. It consists of a fluorescent dye conjugated to Ni^{2+}: nitrilotriacetic acid (NTA) complex. The Ni^{2+} binds specifically to the oligohisitidine domain of the His-tagged fusion protein, allowing specific detection of His-tagged fusion proteins. BenchMark His-tagged Protein Standard acts as a positive control and a protein size marker.

2.7.2.3 Fluorometric enzyme assay for IDS proteins

IDS activity was assayed by the fluorometric method developed in 2001 [Voznyi YV, Keulemans JLM and van Diggelen OP, 2001]. The IDS enzyme assay is composed of two parts: 1) the catalytic reaction of IDS on the substrate desulphation, and 2) the hydrolysis reaction of IDUA on the intermediate to release fluorescent 4-methylumbelliferone (4-Mu). In this study, the IDS activities of three different samples were evaluated: 1)

the lysate of *IDS* transfected COS-7 cells, 2) the total expression products of cell-free systems, and 3) 5-fold diluted plasma samples. For negative control, 0.2 % of pH and heat inactivated BSA was used as testing sample. 10 µL of the testing samples was added to 20 µL substrate (1.25 mM 4-Mu-αIdu-2S). The enzymatic reaction was carried out at 37 °C for 4 hours. Then a mixture containing 40 µL Pi/Ci-buffer and 10 µL LEBT was added to the reaction. After 24-hour incubation at 37 °C, the reaction was terminated with 200 µL stop buffer. Released 4-Mu emits fluorescence at 460 nm upon excitation at 355 nm. The fluorescence of 4-Mu was measured in PerkinElmer 1420 Multilabel Counter VICTOR^3 under the control of Wallac 1420 workstation (PerkinElmer, Inc., Wellesley, USA). Then, the concentration of expression protein products were assayed using Braford method for IDS activity calculations.

2.7.2.4 Luciferase activity assay

The activity of expressed luciferase was assayed by adding 2.5 µL expression products in 50 µL luciferase assay reagent. The luminescence was measured for 10 sec immediately by PerkinElmer 1420 Multilabel Counter VICTOR^3 under the control of Wallac 1420 workstation (PerkinElmer, Inc., Wellesley, USA).

Chapter 3 Results

3.1 Mutational analysis of MPS II and carrier detection

Mutational analysis confirmed that the four MPS II patients have their genetic defects lying in the *IDS* gene. All the fragments containing *IDS* exons were amplified from the patients and sequenced. Two of the families and their female relatives were also screened for mutations. As summarized in Table 3.1, a total of 3 different nucleotide changes were detected in exon 8. These included two nonsense mutations and one missense mutation. One of the two nonsense mutations was previously reported while the other two mutations described here are novel [Jonsson JJ *et al.*, 1995].

A simplified method was used to verify the novel missense mutation. The nucleotide changed by p.L339P (c.1016T>C) created an additional *ScrF I* restriction site within the PCR fragment amplified by *IDS* internal primers F2 and R3 (Figure 3.1). Figure 3.2 shows the partial results of fragment restriction endonuclease analysis. The mutation was not found on screening 100 X-chromosomes in 50 control females. Patient 4 is confirmed to carry a novel mutation in the *IDS* gene. For the novel nonsense mutation, it is unusual to occur as polymorphism in normal individuals. There is no need to perform fragment restriction endonuclease analysis.

Table 3.1 *IDS* mutations identified in this study

Mutation	Nucleotide alteration	Amino acid alteration	Phenotype analysis	Remarks
p.L339P	CTA > CCA	Leucine to Proline	Severe	Novel
p.S369X	TCA > TGA	Serine to terminal codon	Mild	Novel
p.Q389X	CAG > TAG	Glutamine to terminal codon	Severe	

Table 3.1 All the mutations described above are located in exon 8 of *IDS* gene.

Figure 3.1 Potential restriction sites on PCR fragment containing codon 339

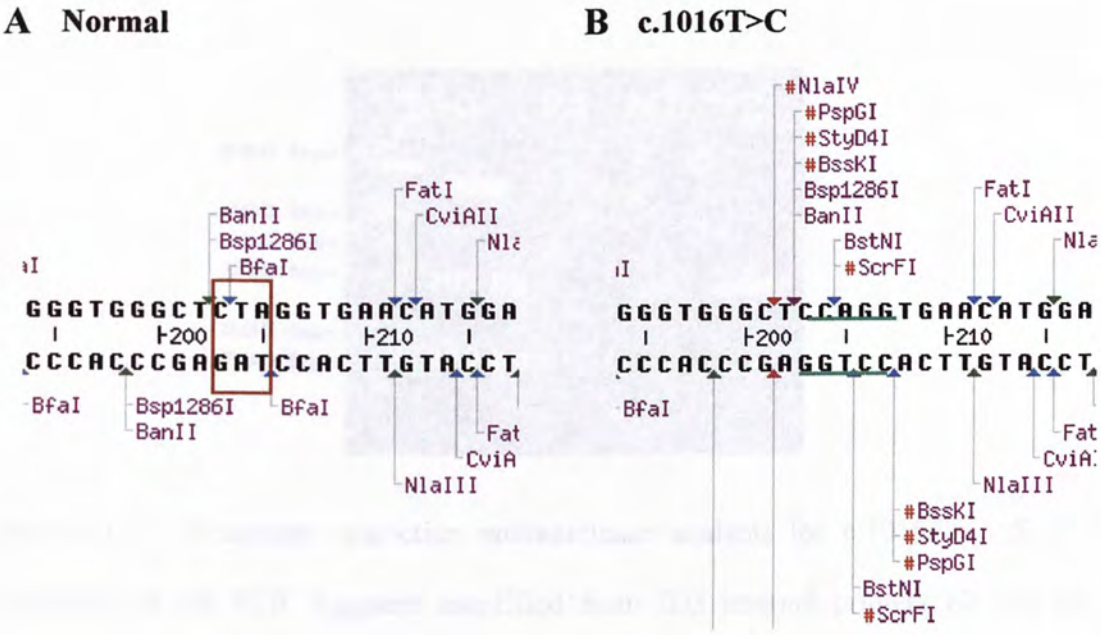


Figure 3.1 Fragment restriction endonuclease analysis for c.1016T>C: *ScrF I* digestion of the PCR fragment amplified from *IDS* internal primers F2 and R3. Codon 339 is boxed and the recognition base pairs (CCAGG) of *ScrF I* is underlined. (A) normal PCR fragment do not have *ScrF I* at codon 339; and (B) c.1016T>C mutant PCR fragment created an additional *ScrF I* restriction site.

Figure 3.2 Partial results of fragment restriction endonuclease analysis for the missense mutation

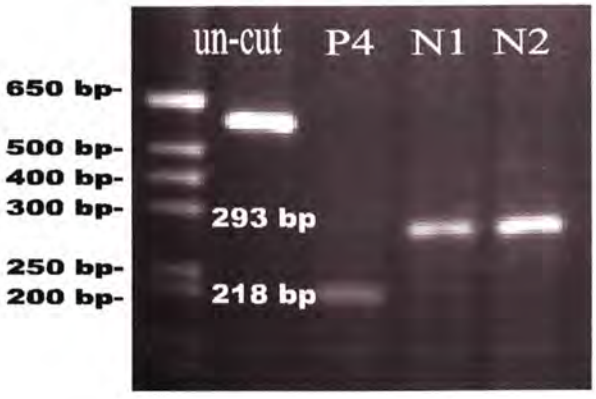


Figure 3.2 Fragment restriction endonuclease analysis for c.1016T>C: *ScrF I* digestion of the PCR fragment amplified from *IDS* internal primers F2 and R3. Un-cut fragment: 622 bp. Normal allele: 293 bp, 183 bp, 87 bp, 39 bp and 20 bp. c.1016T>C mutant allele: 218 bp, 183 bp, 87 bp, 75 bp, 39 bp and 20 bp. It is easier to compare the 293 bp fragment from normal allele with the 218 bp fragment from mutant allele in 3 % agarose gel. Homozygous normal female: 293 bp only. Female carrier: both 293 bp and 218 bp. Affected male: 218 bp only. P4: Patient 4. N1/2: normal individual.

In Family A (Figure 3.3), Patient 1 has the reported mutation identified in the *IDS* exon 8. The codon Gln389 is substituted by a termination codon, i.e., CAG > TAG or p.Q389X. The patient is inherited this mutation from his carrier asymptomatic mother. The screening for his sister and maternal aunt revealed that they are not carriers. The sequencing results are shown in Figure 3.4.

Figure 3.3 Pedigree of Family A

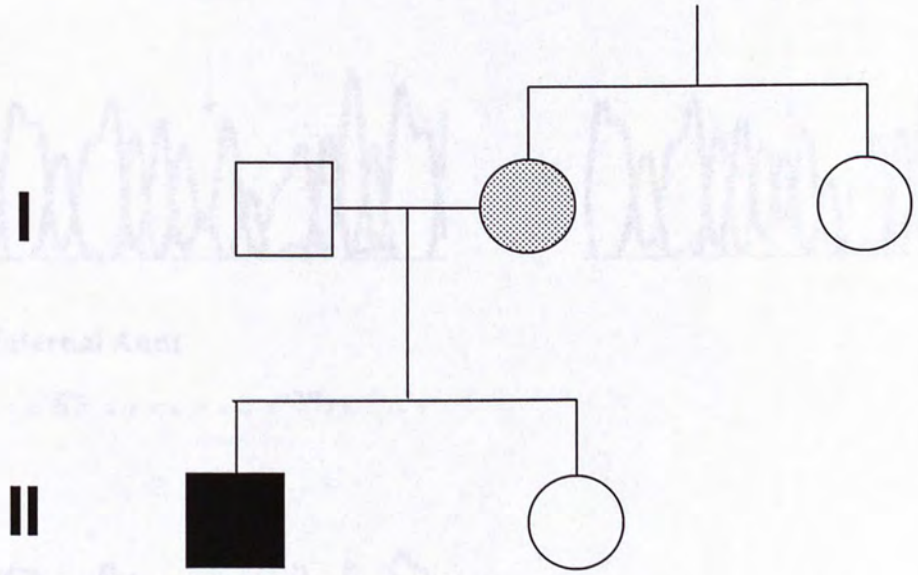


Figure 3.3 Patient 1 inherited the nonsense mutation p.Q389X from his mother. Shaded square represents affected male; dotted circle represents female carrier; and open circle / square means the person is carrying normal *IDS* gene.

Figure 3.4 A reported nonsense mutation, p.Q389X, identified in Family A

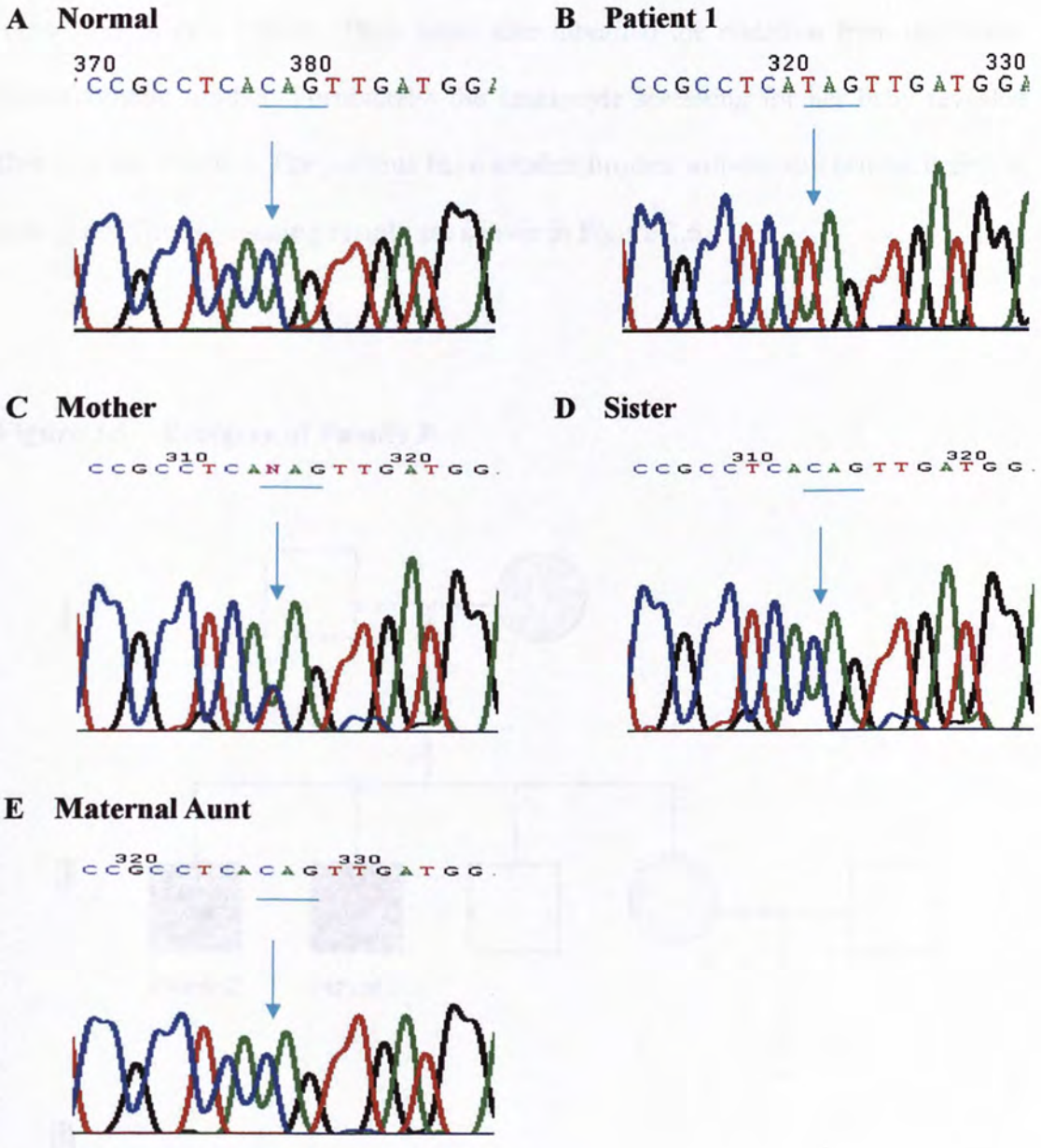


Figure 3.4 Partial sequence electropherograms of *IDS* exon 8 from Family A. The sequence is shown in sense direction. The affected nucleotide is arrowed and the corresponding codon is underlined. Heterozygous base pair is denoted by "N". (A) Normal female; (B) Patient 1 with p.Q389X, i.e., CAG > TAG; (C) the mother; (D) the sister; and (E) the maternal aunt.

In Family B (Figure 3.5), Patient 2 and Patient 3 carry the same novel nonsense mutation in the *IDS* exon 8. The codon Ser369 changed to termination codon, i.e., TCA > TGA or p.S369X. Their sister also inherited the mutation from the carrier asymptomatic mother. Fortunately, the amniocyte screening for her baby revealed that it is not affected. The patients have another brother without any genetic defect in *IDS* gene. The sequencing results are shown in Figure 3.6.

Figure 3.5 Pedigree of Family B

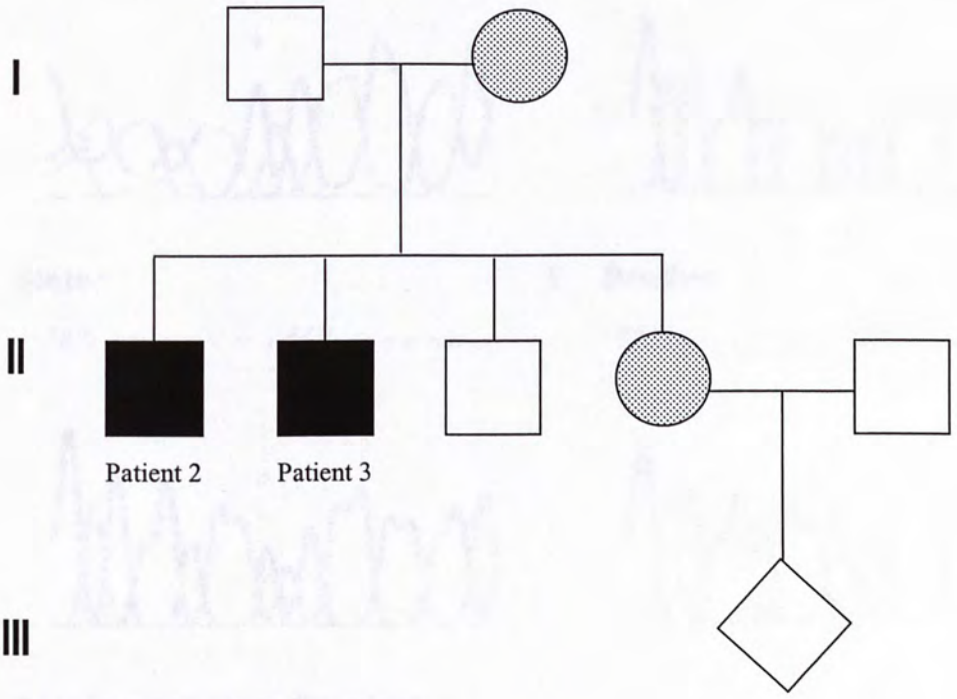


Figure 3.5 Patient 2 and Patient 3 inherited the nonsense mutation p.S369X from their mother. Shaded square represents affected males; dotted circle represents female carriers; and open area means the person is carrying normal *IDS* gene.

Figure 3.6 A novel nonsense mutation, p.S369X, identified in Family B

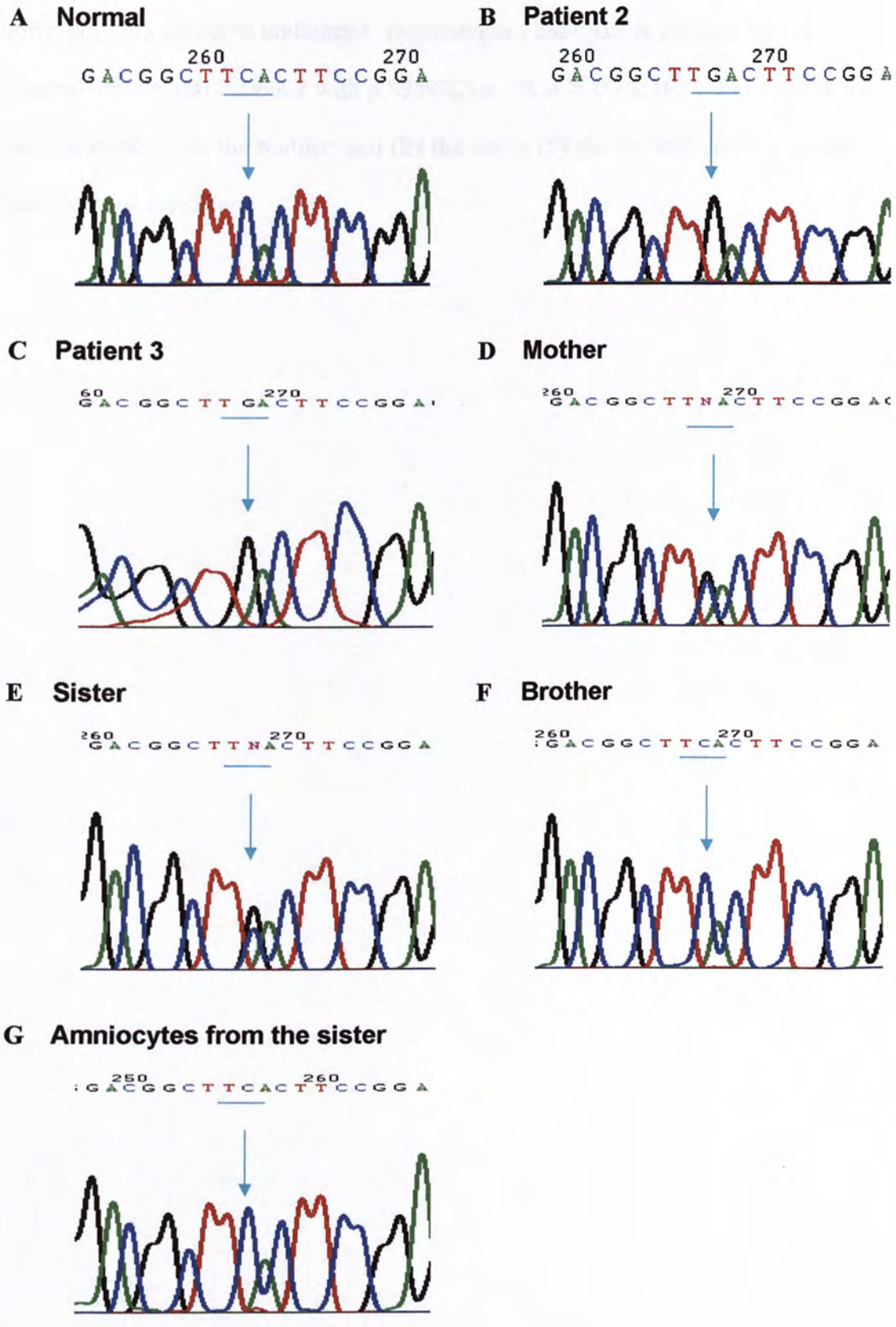


Figure 3.6 Partial sequence electropherograms of *IDS* exon 8 from Family B. The sequence is shown in sense direction. The affected nucleotide is arrowed and the corresponding codon is underlined. Heterozygous base pair is denoted by "N". (A) Normal female; (B) Patient 2 with p.S369X, i.e., TCA > TGA; (C) Patient 3 with the same mutation; (D) the mother; and (E) the sister; (F) the brother; and (G) prenatal diagnosis for the sister..

Patient 4 has a novel missense mutation identified in the *IDS* exon 8, p.L339P. The single nucleotide changed leucine residue to proline, i.e., CTA > CCA or c.1016T> C (Figure 3.7). Since the two premature stop codons described above probably predict protein instability, only the novel missense mutation p.L339P was subjected to further investigations.

Figure 3.7 A novel missense mutation, p.L339P, identified in Patient 4

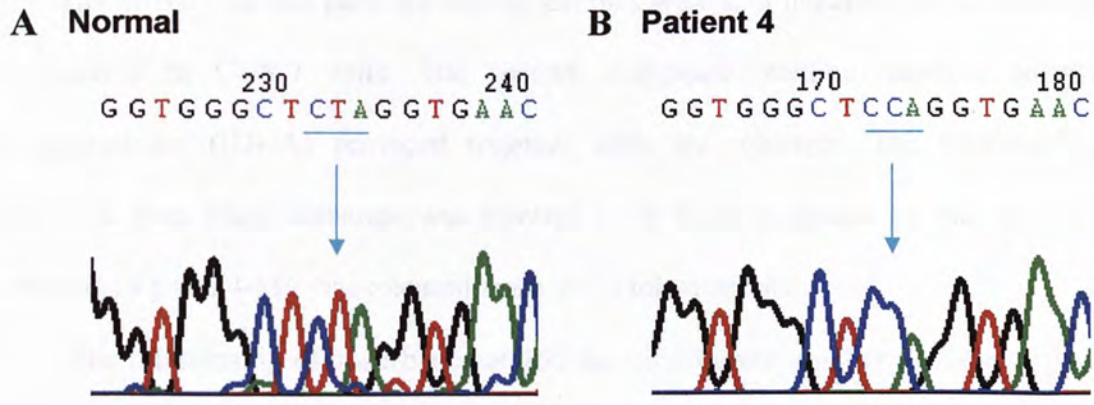


Figure 3.7 Partial sequence electropherograms of *IDS* exon 8 from Patient 4. The sequence is shown in sense direction. The nucleotide changed is arrowed and the corresponding codon is underlined. (A) Normal female; and (B) Patient 4 with p.L339P (c.1016T>C).

3.2 Investigating IDS mutants by transient expression

3.2.1 Fluorometric enzyme assay for measuring IDS activity

The fluorometric enzyme assay was established in 2001 [Voznyi YV *et al.*, 2001]. It involves two steps of enzymatic actions for 1) desulphation of the 4-methylumbelliferone-derived substrates (4-Mu-αIdoA-2S), and 2) hydrolysis of the intermediate to form 4-methylumbelliferone (4-Mu). The first step depends on the IDS activity. In this part, the testing samples were IDS generated from transient expression in COS-7 cells. The second enzymatic reaction depends on the α–iduronidase (IDUA) provided together with the substrate. The fluorescence response from blank-substrate was lowered to 28 % as suggested by the protocol. Around 14 pmol 4-Mu was released from the blank-substrate.

The relationship of absorbance at 460 nm on different number of mole of pure 4-Mu was used as standard curve for subsequence calculations (Figure 3.8 and Figure 3.9). Before the experiment, the assay was validated by measuring plasma IDS activity for 10 normal individuals. The mean plasma IDS activity is 269.44 ± 86.76 nmol/4h/ml. The results lay within the normal range of 167 – 475 nmol/4h/ml IDS activity as suggested by the protocol [Voznyi YV *et al.*, 2001].

Figure 3.8 The linear relationship of fluorometric reaction in IDS enzyme assay

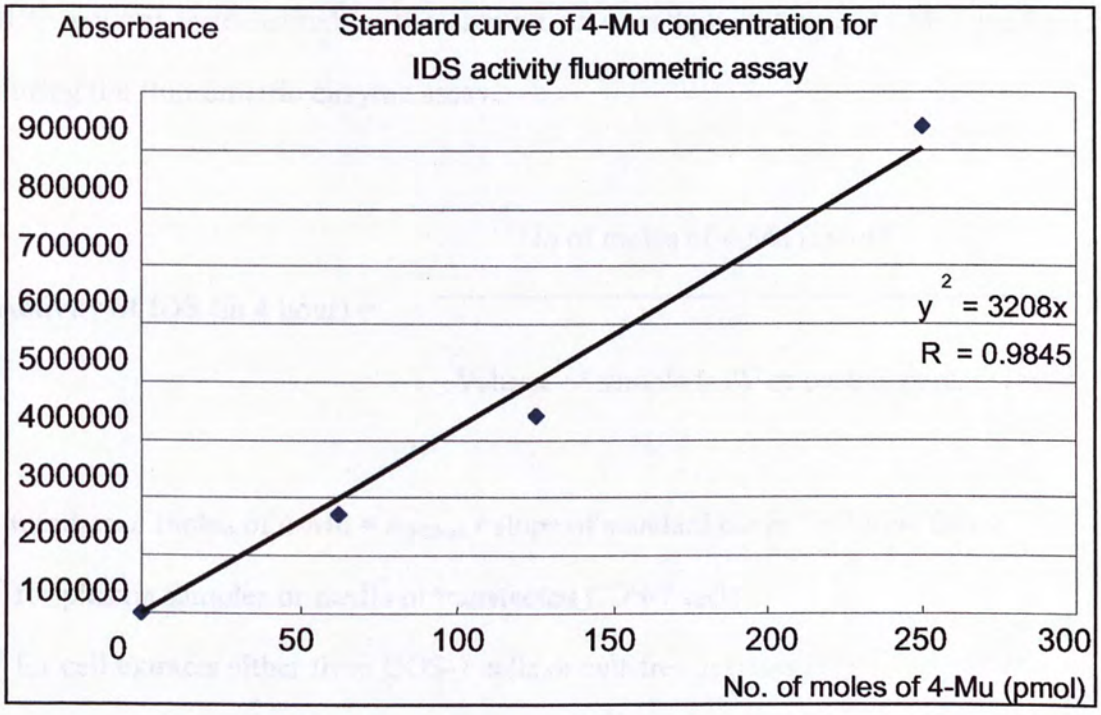


Figure 3.8 The absorbance at 460nm of the enzyme assay shows a linear relationship from 0 - 250 pmol of 4-methylumliferone (4-Mu). 1 pmol of 4-Mu emits 3208 fluorescence response at A_{460nm}.

Figure 3.9 Calculations for IDS activity assay

IDS activity is measured as the amount of 4-methylumliferone (4-Mu) produced during the fluorometric enzyme assay.

$$\text{Activity of IDS (in 4 hour)} = \frac{\text{No of moles of 4-Mu (nmol)}^{a}}{\text{Volume of sample (ml)}^{b} \text{ or protein content (mg)}^{c}}$$

[a] number of moles of 4-Mu = A_{460nm} / slope of standard curve * dilution factor

[b] for plasma samples or media of transfected COS-7 cells

[c] for cell extracts either from COS-7 cells or cell-free systems

For the transient expression products, IDS activity was adjusted with the expression level of co-transfected positive control -- luciferase.

Figure 3.9 IDS activity is expressed as the amount of 4-Mu produced. Only the first step involves IDS enzymatic reaction, so the IDS activity is calculated depending on the first incubation, i.e. 4 hours. The IDS activity is expressed as nmol/4h/mg in cell extracts, nmol/4h/ml in plasma or media of transfected COS-7 cells.

3.2.2 Source of *IDS* gene for transient expression in COS-7 cells

IDS gene was successfully isolated from the normal blood sample with plasma IDS activity 343.1 nmol/4h/ml. Direct DNA sequencing showed that the *IDS* pseudogene was not amplified in the PCR. The wild-type *IDS* cDNA was cloned into the Gateway Entry vector pENTR/D-TOPO and swapped into the mammalian expression vector pcDNA-DEST40. The cDNA sequence is showed in Appendix 3. Both vectors containing wild-type *IDS* cDNA were confirmed by sequencing (Figure 3.10 and Figure 3.11).

The effects of both the mutations at codon 339, p.L339P identified in Patient 4 and p.L339R previously reported in 1998, were investigated [Froissart R *et al.*, 1998]. The two mutant *IDS* cDNAs were produced by site-directed mutagenesis before swapping into the mammalian expression vector. Desired mutations were confirmed by sequencing (Figure 3.10 and Figure 3.11). A modified *IDS* cDNA without 99 base pair at the N-terminal end including the signal sequence was also constructed. This N-terminally truncated *IDS* cDNA was cloned as the same way as the wild-type *IDS* cDNA and confirmed by sequencing (Appendix 3).

Figure 3.10 pENTR/D-*IDS* with desired mutations generated

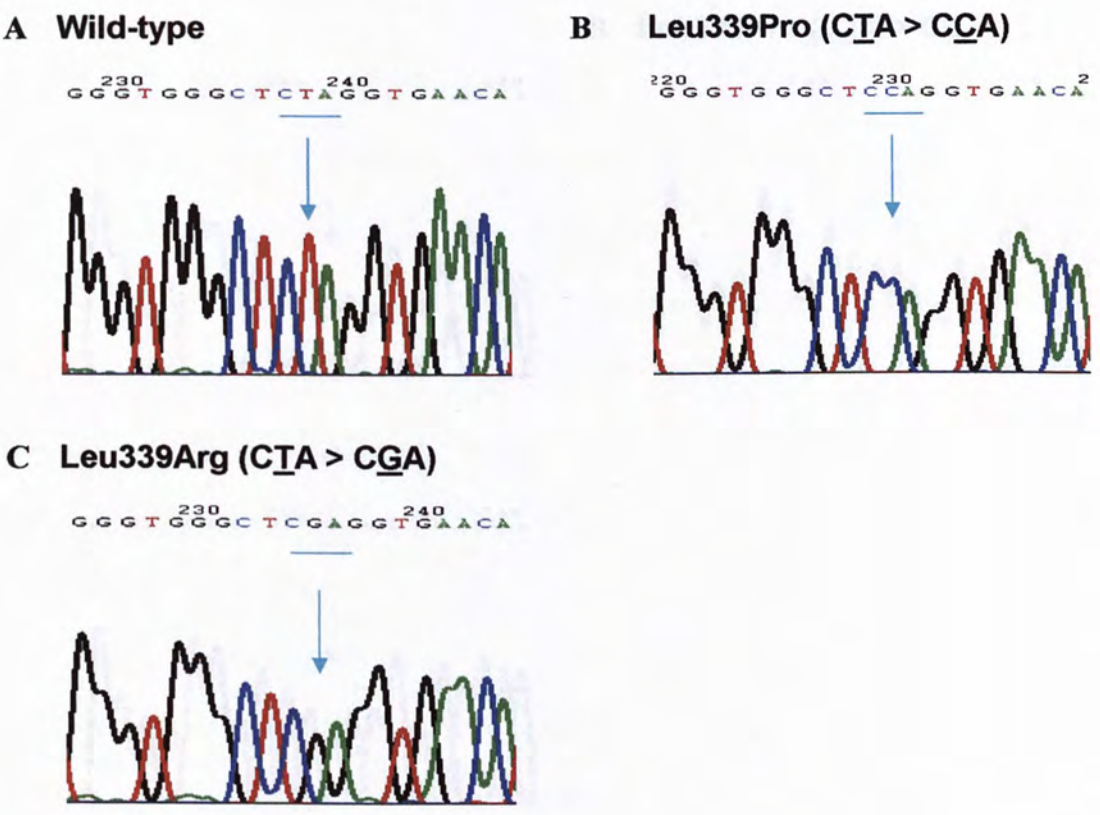


Figure 3.10 Partial sequence electropherograms of the vector pENTR/D-TOPO containing *IDS* cDNA. The plasmids were sequenced by IDS-InterF2 primer and the results are shown in sense direction. The nucleotide changes (arrowed) in the designated codons (underlined), indicated successful site-directed mutagenesis. The whole gene was further sequenced to ensure the absence of undesired mutations. (A) Wild-type *IDS*; (B) Leu339Pro and; (C) Leu339Arg.

Figure 3.11 pcDNA-DEST40-IDS with desired mutations generated

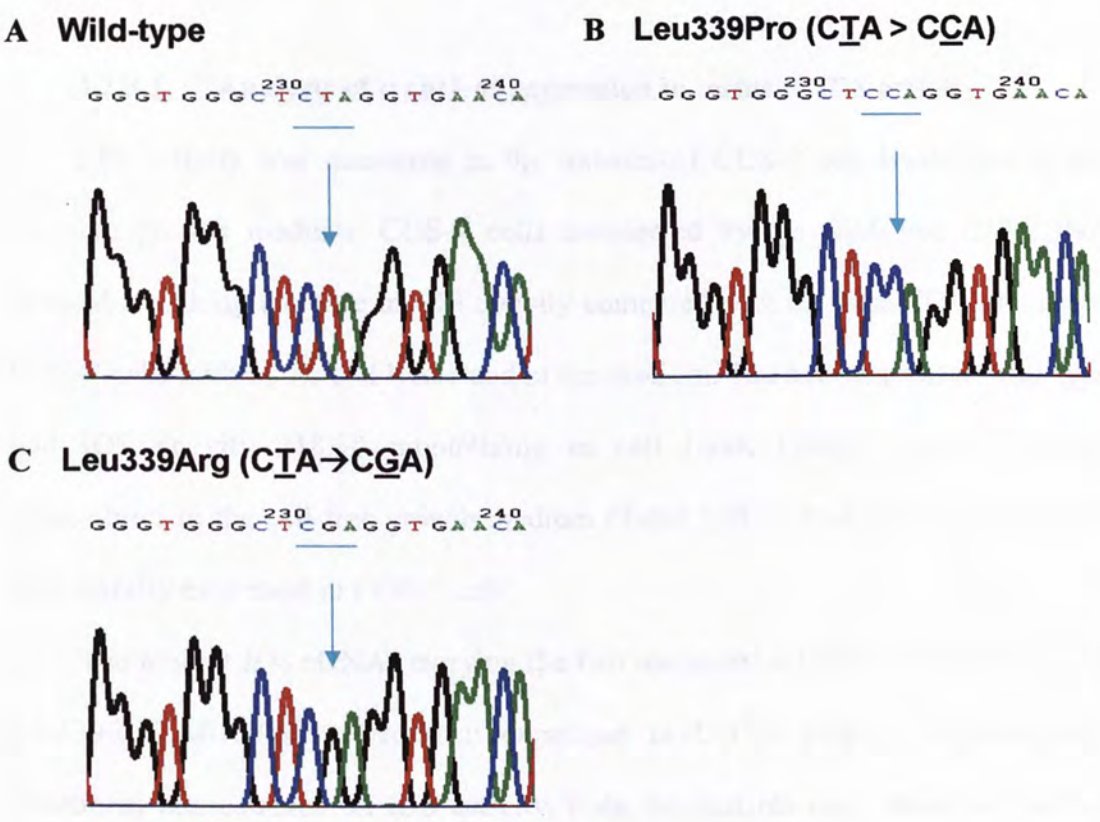


Figure 3.11 Partial sequence electropherograms of the vector pcDNA-DEST40 containing *IDS* cDNA. The plasmids were sequenced by IDS-InterF2 primer and the results are shown in sense direction. The nucleotide changes (arrowed) in the designated codons (underlined), indicated successful swapping of *IDS* gene between Gateway cloning vectors. The whole gene was further sequenced to ensure the absence of undesired mutations. (A) Wild-type *IDS*; (B) Leu339Pro and; (C) Leu339Arg.

3.2.3 *In vitro* expression of IDS and its mutants in COS-7 cells

3.2.3.1 Analysis of transient expression in terms of IDS activity

IDS activity was measured in the transfected COS-7 cell lysate and in the cell-free growth medium. COS-7 cells transfected by the wild-type *IDS* cDNA showed a striking increase in IDS activity compared with the basal IDS activity of COS-7 cells both in the cell lysate and in the medium. The over-expressed wild-type had IDS activity 218.38 nmol/4h/mg in cell lysate (Table 3.2) and 136.68 nmol/4h/ml in the cell-free growth medium (Table 3.3). Active wild-type IDS was successfully expressed in COS-7 cells.

The mutant *IDS* cDNAs carrying the two mutations p.L339P (c.1016T>C) and p.L339R (c.1016T>G) were then expressed in COS-7 cells to evaluate their functional consequence on IDS activity. Both the mutants only remained residual IDS activity less than 2.5 % of the wild-type (Table 3.2). The mutation L339P decreased the IDS activity to 17.86 nmol/4h/mg and the mutation p.L339R had only 20.26 nmol/4h/mg of IDS activity in the cell lysate. The transient expression study demonstrated that both the mutations at *IDS* codon 339, p.L339P and p.L339R, cause significant reduction on IDS activity. The two mutations can reduce the enzymatic activity by introducting protein instability or mRNA instability.

There was no active form of the N-terminally truncated IDS detected in the cell lysate. However, a residual IDS activity more than 50 % of the wild-type was measured in the cell-free medium of transfected COS-7 cells (Table 3.3). Together with the evidence of co-transfected positive control luciferase (data not shown), it can be concluded that all the *IDS* cDNAs were successfully transfected and expressed in the cell-based in vitro expression system.

Table 3.2 Expression of IDS in transfected COS-7 cells.

Transfected *IDS* cDNA	IDS activity (nmol/4h/mg)	Total (percentage of wild-type)
Empty vector (Blank)	15.53	0%
wild-type *IDS*	218.38	100%
p.L339P	17.86	1.15%
p.L339R	20.26	2.33%
N-terminally truncated *IDS*	9.00	0%

Table 3.2 IDS activity was measured in the cell extracts of transfected COS-7 cells. The activities shown here are the means of three independent experiments. The endogenous IDS activity in COS-7 cells was subtracted before the percentage of wild-type was calculated. IDS mutants carrying the mutations p.L339P or p.L339R remained < 2.5 % of wild-type activity. No active form of the N-terminally truncated IDS without the first 33 N-terminal amino acids was detected from the cell lysate.

Table 3.3 Secretion of IDS from transfected COS-7 cells

Transfected *IDS* cDNA	IDS activity (nmol/4h/ml)	Total (percentage of wild-type)
Empty vector (Blank)	39.27	0%
wild-type *IDS*	136.68	100%
p.L339P	20.59	0%
p.L339R	53.14	14.24%
N-terminally truncated *IDS*	98.47	60.77%

Table 3.3 IDS activity was measured in the growth medium of transfected COS-7 cells. The secreted IDS activity in COS-7 cells was subtracted before the percentage of wild-type was calculated. No active form of the mutant IDS carrying the mutation p.L339P was detected in the cell-free media. The N-terminally truncated IDS without signal peptide and the subsequence 8 amino acids conserved > 50 % of wild-type IDS activity.

3.2.3.2 Analysis of *IDS* mRNA stability in COS-7 cells

Secondary structures of partial *IDS* mRNA were predicted by computer programmes (data not shown). The mutation p.L339P created an additional hairpin structure with decreased ΔG -1.04 kcal/mol. Such hairpin structure have the codon 339 CCA forming double-strand mRNA which may induce its degradation. However, this evidence is not strong enough to prove that the mutation p.L339P declined the mRNA stability.

Presence of *IDS* mRNA was confirmed by RT-PCR. The expression level of *IDS* mRNA was normalized by the mRNA of co-transfected *luciferase*. At the same expression level of *luciferase* mRNA, the expression levels of *IDS* mRNA transcripted from different *IDS* constructs were similar (Figure 3.12). This illustrated that the two mutants, p.L339P and p.L339R, have relatively stable mRNA as the wild-type *IDS*. The missense mutations debilitated the IDS enzyme function but not declined the mRNA stability.

3.2.3.3 Analysis of IDS protein stability in COS-7 cells

Normal IDS processing removes parts of the N-terminal sequences and parts of the C-terminal sequences. Fusion protein of IDS carrying GFP or His-tag is not practical. As there are no commercial anti-IDS antibodies for western blotting, only simple protein staining was applied in this study. SDS-PAGE analysis does not show any difference between the COS-7 cells transfected either by wild-type *IDS* or its mutants, or by empty vector (Figure 3.13). The active wild-type IDS was over-expressed without any evidence shown by SDS-PAGE analysis. *Luciferase* was co-transfected as positive control, similarly, no additional protein band or thickened band were observed. SDS-PAGE analysis is not feasible for detecting IDS protein expression in COS-7 cells.

Figure 3.12 RT-PCR of *IDS* mRNA in transfected COS-7 cells

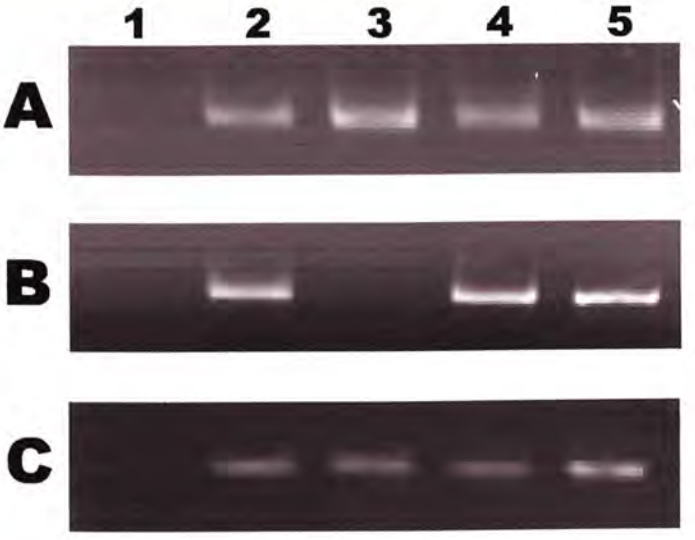


Figure 3.12 RT-PCR of *IDS* cDNA fragments isolated from transfected COS-7 cells. The *IDS* mRNA expression levels are similar between the wild-type *IDS* and its mutants. Panel A: normalization with co-transfected *luciferase* cDNA, 712 bp. Panel B: amplified *IDS* cDNA containing exon 1 to half of exon 9, 1385bp. Panel C: amplified *IDS* cDNA containing only exon 6 to half of exon 9, 621 bp. Lane 1: COS-7 cells transfected by empty vector. Lane 2: COS-7 cells transfected by wild-type *IDS*. Lane 3: COS-7 cells transfected by the N-terminally truncated *IDS* without signal sequence and subsequence 24 bp. Lane 4: COS-7 cells transfected by mutant *IDS* carrying the mutation p.L339P. Lane 5: COS-7 cells transfected by mutant *IDS* carrying the mutation p.L339R.

Figure 3.13 SDS-PAGE analysis for transient expression of IDS and its mutants

Figure 3.13 IDS protein expression in COS-7 cells was shown by SDS-PAGE analysis. *Luciferase* was co-transfected with wild-type IDS or its mutants as the positive control of the expression. Although both wild-type IDS and luciferase were expressed with activities, no additional band or thickened band can be distinguishable between the lanes (expected sizes for IDS and luciferase were 55 kDa and 62 kDa respectively). Lanes 1 – 4: the first independent experiment for the expression of wild-type IDS and its mutants. Lanes 5 – 8: the second independent experiment. Lanes 9 – 12: the third independent experiment. Lanes 4, 8 and 12: total cells extracts from COS-7 cells transfected by empty vector. Lanes 3, 7 and 11: total cells extracts from COS-7 cells transfected by wild-type IDS. Lanes 2, 6 and 10: total cells extracts from COS-7 cells transfected by mutant IDS carrying the mutation p.L339P. Lanes 1, 5 and 9: total cells extracts from COS-7 cells transfected by mutant IDS carrying the mutation p.L339R.

3.3 Cell-free *in vitro* expression for investigating the IDS mutants

3.3.1 The five cell-free systems involved

Both prokaryotic and eukaryotic systems were tested for in vitro expression of wild-type IDS. There were three different cell-free expression methods based on microbial system, plant system and mammalian system. They included two *E.coli*-based systems from Invitrogen and Roche, two wheat germ-based systems from Roche and Promega and a rabbit reticulocytes-based system from Promega. These cell-free expression systems are coupled with transcription and translation in the same reaction tube. The features of each system are summarized in Table 2.4. Successful cell-free expression was confirmed by positive controls: green fluorescence protein (GFP), glucuronidase (GUS) and luciferase.

3.3.2 Source of *IDS* gene for cell-free *in vitro* expression

To express IDS in the cell-free systems, two different T7-driven vectors were used. The wild-type *IDS* and the mutant *IDS* carrying the mutation p.L339P were swapped into another expression vector pEXP1-DEST by Gateway Cloning System. The wild type *IDS* was also cloned into pIVEX-1.3-WG by traditional cloning methods. Successful cloning was confirmed by sequencing (Figure 3.14). These vectors produce N-terminal and C-terminal His-tagged protein respectively for easy detection of the recombinant IDS.

Figure 3.14 pEXP1-DEST-IDS with desired mutations generated

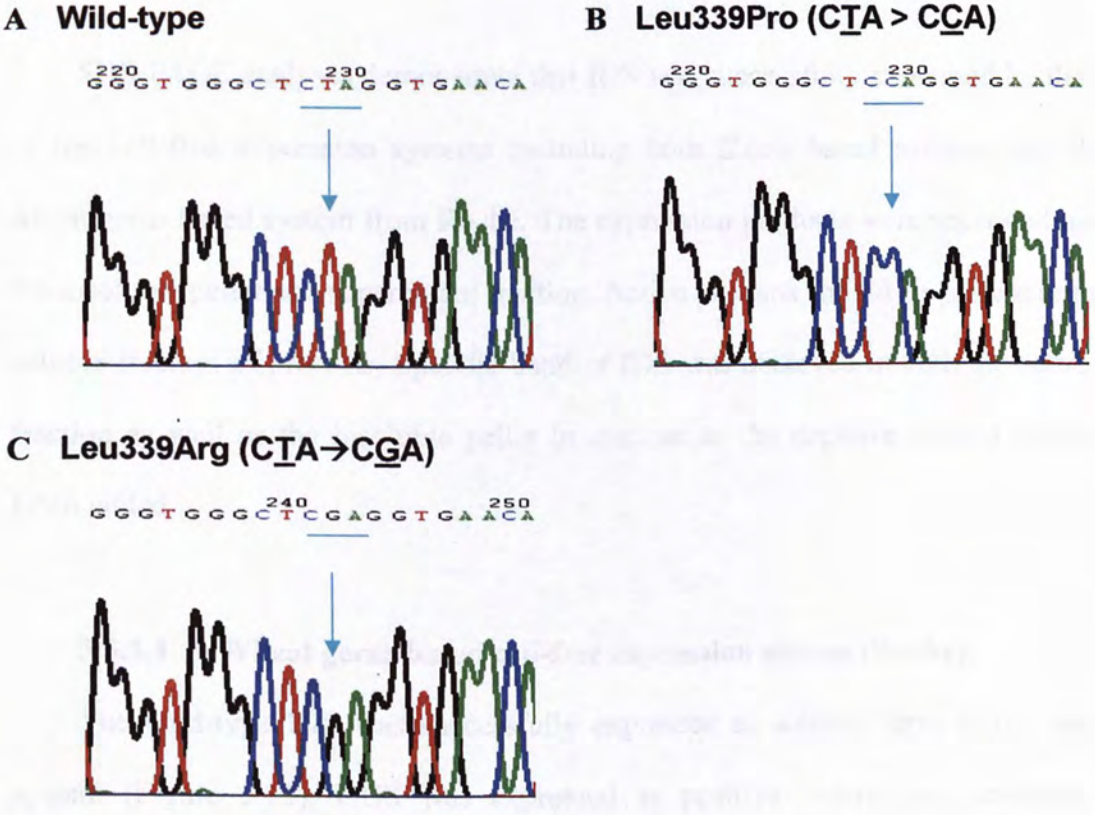


Figure 3.14 Partial sequence electropherograms of the vector pEXP1-DEST containing *IDS* cDNA. The plasmids were sequenced by IDS-InterF2 primer and the results were shown in sense direction. The nucleotide changes (arrowed) in the designated codons (underlined), indicated successful swapping of *IDS* gene between Gateway cloning vectors. The whole gene was further sequenced to ensure the absence of undesired mutations. (A) Wild-type *IDS*; (B) Leu339Pro and; (C) Leu339Arg.

3.3.3 SDS-PAGE analysis of IDS protein stability in cell-free systems

SDS-PAGE analyses demonstrate that IDS was successfully expressed by three of the cell-free expression systems including both *E.coli*–based systems and the wheat germ-based system from Roche. The expression products were separated into the insoluble pellet and supernatant fraction. Active proteins should be present in the soluble fraction of proteins. Specific band of IDS was observed in both the soluble fraction as well as the insoluble pellet in contrast to the negative control without DNA added.

3.3.3.1 Wheat germ-based cell-free expression system (Roche)

The wild-type IDS was successfully expressed as soluble form in the plant system (Figure 3.15). GUS was expressed as positive control and revealed a thickened band with an apparent molecular weight of ~ 70 kDa. However, it was up-shifted to be ~ 85 kDa indicated by the marker. Therefore, the marker Kaleidoscope Prestained Standards may have some deviations in the estimation of protein size. The specific band for wild-type IDS has an over-estimated protein size which must be smaller than 70 kDa. When repeating the SDS-PAGE analysis using another marker BenchMark His-tagged Standard (data not shown), it reveals that specific band of wild-type IDS has molecular weight of 60 kDa.

Although Canine Pancreatic Microsomal Membranes may initiate post-translational processing of proteins, it was recommended not to add into the open system of Promega wheat germ-based cell-free expression system. In this study, it proofed that the microsomal vesicles are also not suitable for adding into Roche wheat germ-based cell-free expression system. The membrane inhibited the overall expression efficiency, no specific band of IDS can be observed (Figure 3.15).

Figure 3.15 SDS-PAGE analysis of wheat germ-based cell-free expression (Roche)

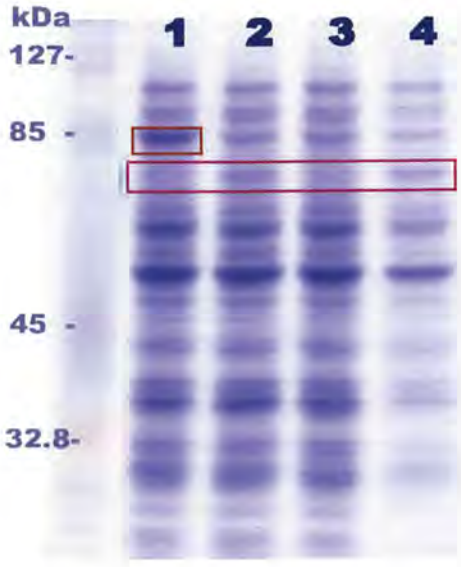


Figure 3.15 Successful synthesis of wild-type IDS was shown by SDS-PAGE analysis. The cell-free expression products were separated into the insoluble pellet and the soluble fraction of proteins. Specific band of wild-type IDS was observed in the supernatant as well as in the pellet when compared with the control. Specific bands of both the wild-type IDS and the positive control GUS were boxed in pink and brown respectively. Only the soluble fraction of proteins was shown. The positive control GUS should be ~ 70 kDa which indicated that the molecular weight was over-estimated by the marker Kaleidoscope Prestained Standards. Lane 1: expression of positive control GUS. Lane 2: expression of wild-type IDS. Lane 3: expression with the presence of canine pancreatic membrane which inhibited the reaction. Lane 4: first washed fluent of purified wild-type IDS in Lane 2. The purification of His-tagged IDS was failed.

3.3.3.2 *E.coli*-based cell-free expression system (Invitrogen)

The wild-type IDS was successfully expressed as insoluble protein in the *E.coli*-based system (Figure 3.16). The absence of the additional band in the supernatant fraction illustrated that the wild-type IDS was likely to be trapped in the inclusion bodies and probably not active. As indicated by the marker, the insoluble IDS had molecular weight similar to the soluble IDS expressed in the plant system, i.e. 60 kDa.

A thickened specific band of the N-terminally truncated IDS was observed in the pellet fraction. Such modified IDS without 33 amino acids was a bit smaller than the wild-type IDS as expected (Figure 3.16). Since there is no proteolytic cleavage of signal peptide by the cell-free system, the N-terminally truncated IDS was expected to fold into active form spontaneously in the presence of RTS GroE Supplement. However, the results illustrated that there was no further processing of C-terminal proteolysis. *E.coli*-based system can only synthesize insoluble IDS either with or without signal peptide.

3.3.3.3 *E.coli*-based cell-free expression system (Roche)

Both the wild-type IDS and the mutant IDS carrying p.L339P were successfully expressed as insoluble proteins in the *E.coli*-based system (Figure 3.17). The absence of the specific band in the supernatant fraction indicated that all the proteins were likely to be trapped in the inclusion bodies and probably not active. Nonetheless, the protein expression level can be compared between the wild-type and the mutant. The precursor protein of mutant IDS with intact signal peptide was relatively stable as the wild-type. The missense mutation p.L339P debilitated the enzyme function but not declined the protein stability of precursor IDS.

RTS GroE Supplement may assist in proper folding of the proteins. However, it

did not work in Roche *E.coli*-based systems. The addition of GroE into the open system of Roche *E.coli*-based system even inhibited the expression of IDS (data not shown). No more specific band of wild-type IDS was observed in any size. Therefore, the GroE supplement and the Canine Microsomal Membranes are not suitable for Roche cell-free expression systems.

3.3.4 InVision His-tag In-gel stain for wild-type IDS and its mutant

The expression of wild-type IDS and its mutant carrying the mutation p.L339P was further confirmed by specific staining His-tagged proteins (Figure 3.18). The recombinant IDS proteins synthesized by the cell-free system were N-terminally His-tagged. Both the specific bands for the wild-type IDS and its mutant were visualized in the total protein fraction with similar intensity. The pellet fraction revealed more apparent results of insoluble wild-type IDS and mutant protein. The two protein bands had similar expression level indicating that both wild-type and mutant protein were relatively stable. Additionally, the absence of specific band of IDS in the soluble fraction is consistent with simple protein staining (Figure 3.17). Figure 3.18 also shows some other His-tagged proteins presented in the Roche *E.coli* extracts both in the supernatant fraction and the total protein fraction. These proteins may be synthesized by Roche for transcription and translation reactions.

For both cell-free expression systems from Promega, no detectable IDS was visualized even by specific staining the His-tagged proteins (data not shown). The expression level in both the Promega systems may be too low.

Figure 3.16 SDS-PAGE analysis of *E.coli*-based cell-free expression (Invitrogen)

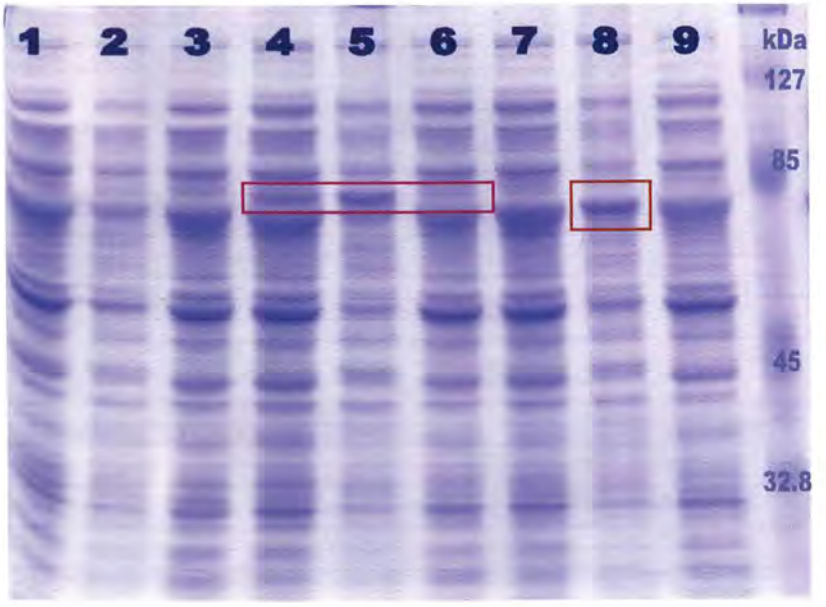


Figure 3.16 Successful synthesis of wild-type IDS was shown by SDS-PAGE analysis. Specific band of the wild-type IDS and the N-terminally truncated IDS was observed in the pellet and the total expression products when compared with the negative control. Wild-type IDS was boxed in pink while the shortened IDS was boxed in brown. Both the IDS proteins have their molecular weight over-estimated. They should be 60 kDa only. Lanes 1 - 3: expression of without DNA template. Lanes 4 - 6: expression of wild-type IDS. Lanes 7 - 9: expression of the N-terminally truncated IDS without 99 bp. Lanes 1, 4 and 7: total expression products. Lanes 2, 5 and 8: insoluble pellet fraction. Lanes 3, 6 and 9: soluble fraction of proteins. The marker was Kaleidoscope Prestained Standards.

Figure 3.17 SDS-PAGE analysis of *E.coli*-based cell-free expression (Roche)

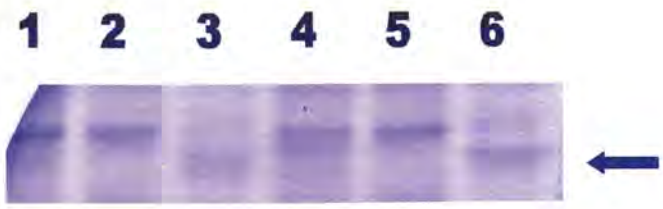


Figure 3.17 Successful synthesis of wild-type IDS and its mutant was shown by SDS-PAGE analysis. Only the samples of wild-type IDS and its mutant were shown. Specific band of the wild-type IDS and the mutant carrying p.L339P was observed in the pellet and the total expression products when compared with the negative control (narrowed). The absence of specific band in the soluble fraction indicated that the cell-free system synthesized insoluble IDS proteins only. The wild-type IDS and its mutant have similar expression level which implied that both the proteins were relatively stable. Lanes 1 - 3: expression of wild-type IDS. Lanes 4 - 6: expression of mutant IDS carrying the mutation p.L339P. Lanes 1 and 4: total expression products. Lanes 2 and 5: soluble fraction of proteins. Lanes 3 and 6: insoluble pellet.

Figure 3.18 His-tagged protein staining of *E.coli*-based cell-free expression (Roche)

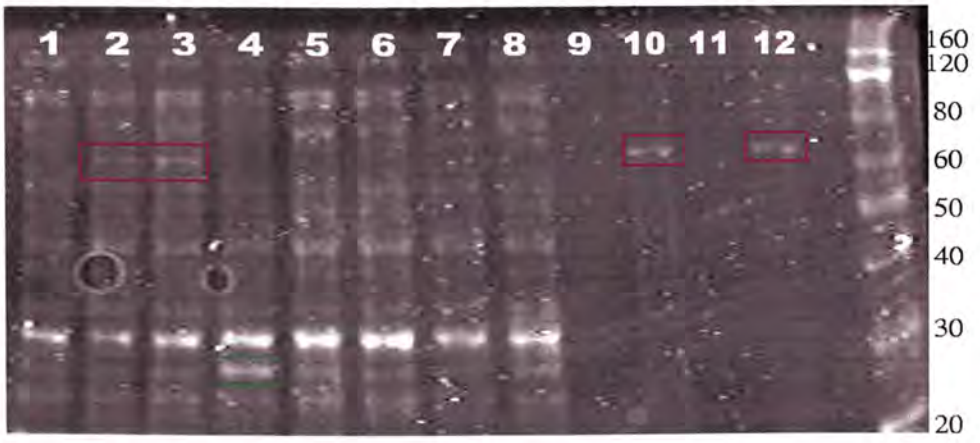


Figure 3.18 Successful synthesis of wild-type IDS and its mutant was further confirmed by His-tag In-gel staining. Specific band of the wild-type IDS and its mutant carrying the mutation p.L339P was observed in the pellet and the total expression products when compared with the negative control. IDS proteins were boxed in pink while the positive control GFP was boxed in green. The absence of specific band in the soluble fraction indicated that the cell-free system synthesized insoluble IDS proteins only. The wild-type IDS and its mutant have similar expression level which implied that both the proteins were relatively stable. Lanes 1 - 4: total expression products. Lanes 5 - 8: soluble fraction of proteins. Lanes 10 and 12: insoluble pellet. Lanes 1 and 5: expression without DNA template. Lanes 2, 6 and 10: expression of wild-type IDS. Lanes 3, 7 and 12: expression of mutant IDS carrying p.L339P. Lanes 4 and 8: expression of positive control GFP. Lanes 9 and 11: no sample loaded. The marker was BenchMark His-tagged Protein Standard.

3.3.5 Analysis of IDS activity in cell-free expression systems

The overall performance of each cell-free expression systems was summarized in Table 3.4. The IDS activity assay was tested on the total expression products. All the IDS protein synthesized in either of the cell-free systems was inactive when compared with the negative control (Figure 3.19). Although SDS-PAGE analysis indicated successful protein expression by the cell-free systems, most of the recombinant IDS was insoluble. The wild-type IDS expressed by Roche wheat germ-based system was not active even in soluble form. Additionally, the assay results confirmed that no active IDS was synthesized by the two Promega cell-free expression systems. For the mammalian system, negative results were reproduced by using the expression vector pcDNA-DEST40 containing *IDS* cDNA (data not shown). Since the cell-free systems failed to produce active wild-type IDS, the mutant was not subjected to further activity assay.

As shown by SDS-PAGE analysis (Figure 3.15), purification of His-tagged protein was failed. The His-tagged wild-type IDS synthesized by the Roche wheat germ-based system was tried to isolate from the total expression products. However, there was a striking amount of wild-type IDS rinsed away in the first wash (Figure 3.15). The purified products and the fluent from the washing steps were not subjected to further activity assay.

Table 3.4 Summary of IDS expression in COS-7 cells and in cell-free systems

	COS-7	E.Coli (Invitrogen)	E.Coli (Roche)	Wheat germ (Roche)	Wheat germ (Promega)	Reticulocyte (Promega)
Expression of positive control	+ Luci	+ Luci + GFP	+ Luci + GFP	+ GUS	+ Luci	+ Luci
IDS protein expression	+	Insoluble precursor	Insoluble precursor	soluble & insoluble precursors	not visualized	not visualized
IDS catalytic activity	+	not detectable	not detectable	not detectable	not detectable	not detectable

Table 3.4 The overall performance of all the expression systems. Only the cell-based method using COS-7 cells produce active IDS protein. None of the cell-free system expressed active wild-type IDS although some of the systems showed good protein expression of IDS. "+" means positive results; "+ Luci" means successful expression of active luciferase; "+ GFP" means successful expression of GFP showing fluorescence under UV light; "+ GUS" means successful expression of GUS with evidence shown by SDS-PAGE.

Figure 3.19 Analysis of cell-free expression in terms of IDS activity

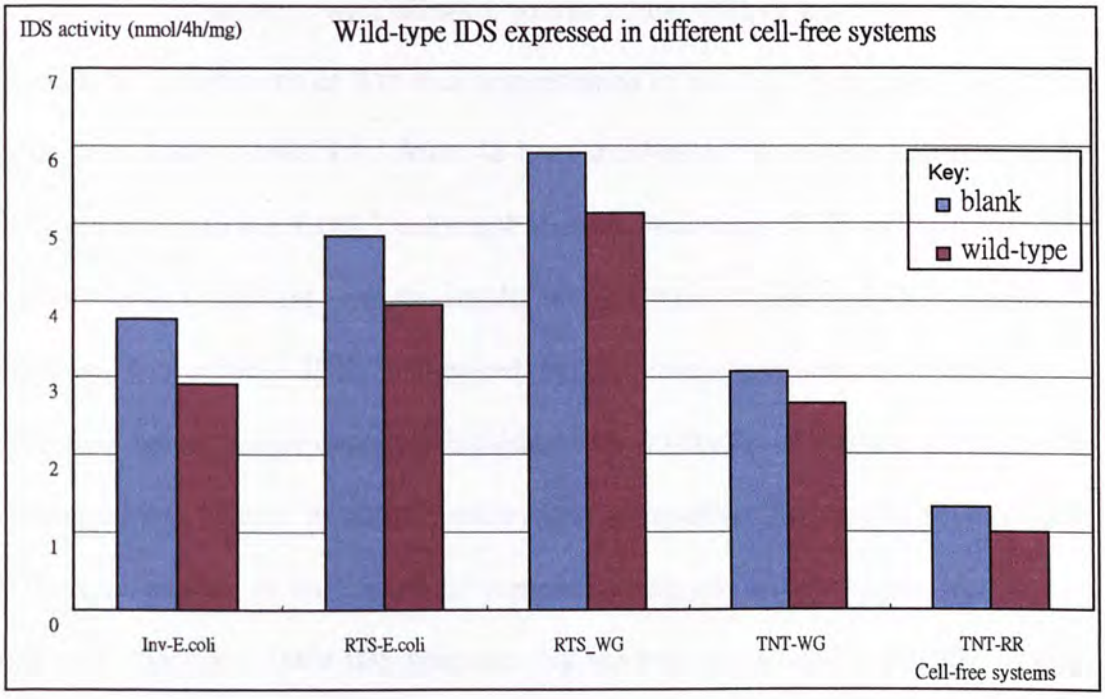


Figure 3.19 Wild-type IDS synthesized by any of the cell-free systems was not active. Each expression was repeated at least three times. All the absorbance in the activity assay lay in the linear range, indicating that the activity assay was not yet saturated and the results were reliable. Successful protein expression of the different cell-free systems was also confirmed by active luciferase co-expressed as positive control. Blank: expression reagents without DNA template. Wild-type: expression of wild-type IDS. Inv-E.coli: *E.coli*-based expression system Invitrogen. RTS-E.coli: *E.coli*-based expression system from Roche. RTS-WG: wheat germ-based expression system from Roche. TNT-WG: wheat germ-based expression system from Promega. TNT-RR: rabbit reticulocytes-based expression system from Promega.

3.3.6 Analysis of the cellular uptake of IDS

The IDS protein was allowed to enter the COS-7 cells by endocytosis. Successful endocytosis of IDS was demonstrated by active IDS expressed in COS-7 cells previously (Table 3.5). After 48-hour incubation, the N-terminally truncated IDS entered into the COS-7 cells and showed more than 50 % of wild-type IDS activity. It is consistant with the results obtained from transfected COS-7 cells. In contrast, the soluble IDS synthesized by the wheat germ-based Roche system remained inactive after endocytosis (Table 3.6). It is believed that precursor may be processed and resume its activity after entering the cells. The results of the soluble IDS were similar to the results of expression reagents without DNA template. It indicated that the soluble IDS precursor did not enter the COS-7 cells after 48-hour incubation. Thus, the insoluble IDS synthesized by the other cell-free systems was not subjected to further endocytosis study in COS-7 cells.

Table 3.5 Endocytosis study of active IDS into un-transfected COS-7 cells

		Incubation	IDS activity
Donor	Medium of wild-type *IDS* transfected COS-7 cells		136.68[a]
Recipient	Un-transfected COS-7 cells	0 hour	0.98[b]
	Un-transfected COS-7 cells	48 hours	1.26[b]
Donor	Medium of N-terminally truncated *IDS* transfected COS-7 cells		98.47[a]
Recipient	Un-transfected COS-7 cells	0 hour	0.54[b]
	Un-transfected COS-7 cells	48 hours	0.98[b]

[a] IDS activity expressed in nmol/4h/ml; [b] IDS activity expressed in nmol/4h/mg

Table 3.5 Cellular uptake of active IDS into un-transfected COS-7 cells was successful. The N-terminally truncated IDS remained > 50 % of wild-type activity after endocytosis. The result was consistent with transient expression. The activity differences before and after endocytosis were similar for both the two types of IDS.

Table 3.6 Endocytosis study of the total expression products from cell-free system

		Incubation	IDS activity (nmol/4h/mg)
Donor	Negative control from wheat germ-based system (Roche)		5.91
Recipent	Un-transfected COS-7 cells	0 hour	3.90
	Un-transfected COS-7 cells	48 hours	0.65
Donor	Wild-type IDS from wheat germ-based system (Roche)		5.12
Recipent	Un-transfected COS-7 cells	0 hour	5.19
	Un-transfected COS-7 cells	48 hours	0.58

Table 3.6 Endocytosis of the total expression products containing soluble IDS synthesized by the Roche wheat germ-based system. The activity of wild-type IDS was lower than the negative control after endocytosis. No active form of the soluble IDS was generated after endocytosis. Insoluble IDS expressed by the other cell-free expression systems was not subjected to further endocytosis study.

Chapter 4 Discussions

4.1 Mutational analysis

4.1.1 Heterogeneity of *IDS* mutations

MPS II is a rare genetic disease. With the two novel mutations described in this study, a total of 321 mutations have already been identified in the *IDS* locus. It indicates a high allelic heterogeneity in the *IDS* gene and which is responsible for the broad spectrum of clinical severity observed in MPS II patients [Hopwood JJ *et al.*, 1993]. As mentioned in Chapter 1.2.3.2, exons 3, 5, 8 and 9 are the hot spots of point mutation. Of which, 12.3 % of reported point mutations, predicting amino acid substitution and translational termination, were identified in exon 8. In this study, all the three mutations are also located in exon 8. It further demonstrates that *IDS* exon 8 is the hot spot of point mutations in Chinese population [Liu S *et al.*, 2002].

4.1.2 Role of molecular diagnosis for MPS II

Mucopolysaccharidoses are suspected when elevated GAG is detected in the urine samples. However, MPS I and MPS II share the same biochemical markers in this preliminary test. Since IDUA and IDS are involved to complete the first step of GAG degradations (Figure 1.1), both types of patient excrete the common substrates in massive amount. In order to distinguish between MPS I and MPS II when elevated heparin sulfate and dermatan sulfate are detected, further investigations

must be done. MPS II can be confirmed by demonstrating the IDS deficiency in the patients plasma or cultured fibroblasts. However, it may take days to weeks to confirm the results. A more efficient method is necessary, especially for prenatal diagnosis.

Majority of LSDs only require low enzymatic activity to maintain normal phenotypes although the minimum requirement of IDS is unknown for MPS II patients [Thomas GH, 1994]. It is also difficult to differentiate the clinical severity of MPS II based on residual IDS activity [Parkinson EJ *et al.*, 2004; Kato T *et al.*, 2005]. On the other hand, recent studies established some genotype-phenotype correlations by mutational analysis as described in Chapter 1.2.3.3 [Millat G *et al.*, 1998; Dierks T *et al.*, 1999; Moreira da Silva I *et al.*, 2001; Kato T *et al.*, 2005]. Therefore, direct DNA sequencing on *IDS* gene serves as a better diagnostic method for MPS II. It provides definitive results with high specificity and the automated procedures speed up the screening process within one-day. Having the mutation identified in Patient 2 and Patient 3, prenatal diagnosis for their sister was performed efficiently in this study.

In addition, mutational analysis can replace enzyme assay for accurate carrier detection [Froissart R *et al.*, 1998; Keulemans JLM *et al.*, 2002]. Skewed X-chromosome inactivation may occur in some carriers and lead to lower enzymatic activity as in the patients [Lyon MF, 1961; Cudry S *et al.*, 2000]. On the other hand, the identified mutations found in the patients help effective screening for the female carriers as in this study. Three MPS II female carriers including the patients' mothers were confirmed by direct DNA sequencing. The results showed that the two nonsense mutations are not de novo but inherited from the patients' mothers.

4.1.3 Two novel mutations and one reported mutation were identified

4.1.3.1 A novel nonsense mutation: Ser369term

Patient 2 and Patient 3 are carrying the same mutation p.S369X in the *IDS* gene which is responsible for mild form of MPS II. Since the two patients are brothers, the frequency of p.S369X only counts once as in Table 1.2. The nonsense mutation predicts termination of translation from Ser369 residue and leads to C-terminally truncated protein without 181 amino acids. It is believed that Ser369 residue is not the second proteolytic cleavage site for normal C-terminal processing otherwise the nonsense mutation would not lead to MPS II [Wilson PJ *et al.*, 1990; Millat G *et al.*, 1997a]. Kato's group predicted that the last putative active site is located at Lys347 residue [Kato T *et al.*, 2005]. Therefore, the mutant protein carrying p.S369X should have all the catalytic domain and subdomains preserved for residual enzymatic activity. IDS deficiency may be due to protein instability. Including this novel mutation, there are 33 different nonsense mutations reported. And they account for nearly 10 % of total mutations identified in the *IDS* locus.

4.1.3.2 A reported nonsense mutation: Gln389term

Patient 1 carried the mutation p.Q389X in the *IDS* gene which is responsible for clinically severe MPS II. This nonsense mutation was first reported in America [Jonsson JJ *et al.*, 1995]. The translation is likely to terminate at Gln389 residue leading to the loss of 161 amino acids at the C-terminal end. The un-translated sequence does not contain any putative active site [Kato T *et al.*, 2005]. p.Q389X should result in shortened polypeptides with residual enzymatic activity similar to p.S369X described above. Interestingly, with the addition of 20 amino acids following residue 369, the disease becomes a severe form. Together with the mild

forms having nonsense mutations in exon 1, such as W12X, it indicates that structural analysis is not enough to explain the pathological lesion of nonsense mutation in *IDS* gene [Filocamo M *et al.*, 2001; Kato T *et al.*, 2005].

4.1.3.3 A novel missense mutation: Leu339Pro

In the severely affected Patient 4, a novel missense mutation p.L339P was identified. The point mutation predicts substitution of Leu339 residue by proline. The restriction analysis revealed the absence of c.1016T>C in at least 50 normal control individuals. However, it is a non-conservative change when comparing at least ten human sulfatases (Appendix 1). And Leu339 residue does not belong to the active C-X-P-S-R motif and it is not a potential N-glycosylation site [Millat G *et al.*, 1997a]. Indeed, it is not enough to support the view that single base pair change is the pathological lesion for clinically severe MPS II. Previous transient expression studies provided major evidence for the causality of most missense mutations [Bonuccelli G *et al.*, 2001; Chang JH *et al.*, 2005]. Therefore, *in vitro* expression of the mutant gene was involved in this study for verifying the enzyme defect.

4.2 Expression studies of the IDS mutants

Besides the novel mutation p.L339P, there is another missense mutation reported at the same codon 339, i.e. p.L339R [Froissart R *et al.*, 1998]. Although the deleterious effect of p.L339R was not verified by expression study previously, it was reported to be responsible for clinically severe MPS II. Therefore, it is worth to investigate both the missense mutations at codon 339 for their effects of amino acids alterations and the functional importance of Leu339 residue. p.L339P and p.L339R were studied through site-directed mutagenesis and *in vitro* expression. Transient expression in COS-7 cells demonstrated that both of the missense mutations caused significant consequences on IDS catalytic activity but did not affect the mRNA stability. And the *in vitro* expression using cell-free systems demonstrated the mutant proteins carrying p.L339P were relatively stable as wild-type *IDS*.

4.2.1 Analysis of transient expression in COS-7 cells

To evaluate the functional consequence on enzyme activity of IDS mutants, it is common to carry out expression studies in COS-7 cells [Villani GR *et al.*, 2000; Bonuccelli G *et al.*, 2001; Ricci V *et al.*, 2003; Chang JH *et al.*, 2005]. The mammalian cells serve as a good tool for which provide all the necessary post-translational modifications including FGly formation, glycosylation and proteolytic processing. Additionally, human IDS maturation steps in COS-7 cells are identical in human fibroblasts and lymphoblastoid cell lines [Wilson PJ *et al.*, 1990; Millat G *et al.*, 1997a, 1997b, 1998; Cudry S *et al.*, 2000]. In COS-7 cells, the two precursor forms of IDS, 76 kDa and 90 kDa are processed through a series of

intermediates to two mature polypeptides, 55 kDa and 45 kDa. The immunoprecipitation of ^{35}S-labeled IDS revealed that all the precursors were processed 48-hour after transfection leaving mature forms in the cells [Millat G *et al.*, 1997b]. Therefore, all the harvested IDS were fully processed in this study.

A positive control vector containing firefly *luciferase* was synthesized for co-transfection with *IDS* constructs. Similar to IDS, luciferase polypeptides also consist of 550 amino acids but no modification required. With the use of CMV promoter and the use of lipofection procedures, it is believed that both luciferase proteins and IDS proteins have the same transfection efficiency and expression efficiency. Luciferase and wild-type IDS were successfully expressed with positive activities in the COS-7 cells. However, no specific band of either luciferase or IDS was observed by simple protein staining. The expression level of both the products may be very low. Therefore, simple protein staining is not feasible for direct visualization of IDS in transfected COS-7 cell lysate as in normal human fibroblast lysate [Millat G *et al.*, 1997b].

There are some other limitations when using COS-7 cells. One of which is the basal IDS activity in the mammalian cells. Other studies have used enzyme-deficient fibroblasts for transient expression study [Sukegawa K *et al.*, 1995; Tomanin R *et al.*, 2002]. Nevertheless, the wild-type IDS was successfully over-expressed resulted in a 13-fold increase in catalytic activity and only the relative activities were required for comparison in this study. The other drawbacks are the low transfection efficiency and low expression efficiency of the mammalian cells. In this study, the use of cationic lipid transfectants provides higher transfection efficiency for the COS-7 cells [Ciccarone *et al.*, 1999; Chang JH *et al.*, 2005]. And the expression efficiency was maximized by using expression vectors containing CMV promoter. Additionally, both the wild-type and mutant constructs were subjected to the same expression

conditions. Therefore, neither transfection efficiency nor expression efficiency is a major problem for investigating the causality of the two missense mutations.

4.2.1.1 Stability of mutant mRNA

The expression constructs contained the cDNA of either wild-type *IDS* or the mutants *IDS*. So there is no mRNA splicing required after transcription in COS-7 cells. The total mRNA level of *IDS* was estimated by RT-PCR. The N-terminally truncated *IDS* without the 33 amino acids at the N-terminal end served as a second negative control. This construct also resulted in no specific band for the amplicon of *IDS* from exon 1 to exon 9 indicating that the basal *IDS* mRNA expression in COS-7 cells is negligible (Figure 3.13B). When compared with the wild-type *IDS* mRNA, both the mutants have similar mRNA expression level and their mRNA is relatively stable in the COS-7 cells (Figure 3.13B & C). Therefore, both the single base pair changes, CTA > CCA and CTA > CGA, at *IDS* codon 339 did not decline the mRNA stability in general.

4.2.1.2 IDS catalytic activity

IDS activity is routinely assayed using radiolabelled disaccharides substrate derived from heparin [Guglielmo RD *et al.*, 2000; Bonuccelli G *et al.*, 2001; Hong YT *et al.*, 2003]. Obviously, 4-methylumbelliferyl derived substrates used in the fluorometric method is more convenient and less cumbersome than the conventional assay [Voznyi YV, Keulemans JLM and van Diggelen OP, 2001]. After the development of the fluorometric assay for IDS in 2001, prenatal diagnosis for MPS II began to apply this method to screen the cultured CV-cells [Keulemans JLM *et al.*, 2002]. Due to its sensitivity and specificity, the fluorometric assay is also used in various expression studies as well as in the present study [Tomanin R *et al.*, 2002;

Ricci V *et al.*, 2003; Chang JH *et al.*, 2005]. Nowadays, fluorogenic substrates are widely used for most other MPS diagnosis, e.g. MPS III A/C, MPS IV A, and for other enzymes like α–L-fucosidase and β–galactosidase [Cuer M *et al.*, 2000; McGuire JBJ *et al.*, 2002].

The transient expression study demonstrated the functional consequence of both missense mutations p.L339P (novel) and p.L339R. Both the amino acids alterations at Leu339 residue cause significant reduction on IDS activity, only less than 2.5 % of normal level, which illustrated the deleterious nature of both the mutations (Table 3.2). The results confirmed that p.L339P identified in Patient 4 is a novel disease-causing missense mutation. As mentioned in Chapter 1.1.4.1, the clinical severity does not depend on the deficient IDS enzyme activity, ranging from no detectable activity to residual activity [Li P, Bellows AB and Thompson JN 1999; Parkinson EJ *et al.*, 2004; Kato T *et al.*, 2005]. And the minimum active IDS level for normal phenotypes remains unknown [Thomas GH, 1994]. Nevertheless, both p.L339P and p.L339R were identified in severely affected patients. The results confirmed that both the point mutations at codon 339 are correlated with clinically severe MPS II.

4.2.2 Analysis of mutant stability by cell-free expression systems

To date, there are no commercial anti-IDS antibodies available for Western blotting [Cudry S *et al.*, 2000]. And Figure 3.12 shows that IDS synthesized in COS-7 cells was not visualized by simple protein staining. Alternative methods were used to synthesize His-tagged recombinant IDS for specific staining. If the substitution of proline at Leu339 residue really changes the conformation of IDS

polypeptide and affects its stability, the precursor would be easily degraded due to improper folding [Tikkanen R *et al.*, 1995]. In the study, detectable amount of IDS precursors were successfully synthesized by the cell-free expression systems. SDS-PAGE analysis, using simple protein staining and His-tagged protein staining, revealed that both the wild-type IDS and mutant precursors were quite stable. Therefore, the novel missense mutation p.L339P does not decline the protein stability.

4.2.3 Structural analysis of amino acids alterations

Structural characteristics of *IDS* mutations were investigated in several groups although there is no official 3D structure of mature IDS published to date [Kim CH *et al.*, 2003; Parkinson-Lawrence E *et al.*, 2005; Kato T *et al.*, 2005]. A 3D structure of the enzyme is useful for analysis of genotype/phenotype correlations in some diseases such as MPS IVA [Sukegawa K *et al.*, 2000; Kato T *et al.*, 2005]. It could clearly indicate molecular effects of the mutations. Generally, great structural change in the mutant IDS is responsible for clincially severe MPS II [Parkinson-Lawrence E *et al.*, 2005]. Missense mutations found in the severe MPS II patients would have direct interactions with the active site residues [Kato T *et al.*, 2005]. And for those mutations found in the mild form of MPS II were located in the peripheral region of IDS.

In a non-conservative residue, which does not conserved in other human sulfatase and does not belong to the sulfatases C-X-P-S-R motif, different amino acids alterations at the particular residue may lead to different clinical consequence. For instance, patients inherited the missense mutation p.R95T were mildly affected

with normal intelligence while patients inherited p.R95G were classified as intermediate with a little mental retardation [Goldenfum SL *et al.*, 1996; Dierks T *et al.*, 1999; Moreira da Silva I *et al.*, 2001]. In addition, IDS deficiency caused by cyclic proline substitution at different leucine residues also resulted in different clinical severity. For instance, p.L182P leaded to intermediate form of MPS II with subnormal intelligence but p.L259P leaded to severe phenotypes with mental retardation [Isogai K *et al.*, 1998; Kim CH *et al.*, 2003].

4.2.3.1 p.L339P causes conformational change

In this study, the novel mutation p.L339P have the hydrophobic Leu339 residue replaced by the equally non-polar but ambivalent cyclic proline. Leucine belongs to the aliphatic R groups which generally buried inside the folded protein molecules. However, proline can exist as *trans-* or *cis-* isomers and is responsible for protein folding. The low residual activity of the IDS mutant suggested that the single base pair change probably affects the proper folding of the polypeptides. In this study, IDS enzymatic activity was also measured in the medium of transfected COS-7 cells because IDS proteins were released into the cell-free medium spontaneously from the senescent cells. Interestingly, there was no detectable IDS activity of the mutants carrying p.L339P. It implied that conformational changes caused by p.L339P may also affect the protein leaving the cells.

4.2.3.2 p.L339R changes overall charge balance

For the reported mutation identified at the same codon, i.e. p.L339R, the non-polar and hydrophobic Leu339 residue was substituted by polar and very hydrophilic arginine. As a cation, arginine may alter the overall charge of the IDS polypeptides. The deleterious effect of p.L339R may due to the disturbance of

overall charge balance of the mutant protein and cause deficient IDS. Therefore, the substitutions of different amino acids at Leu339 residue may result in different degree of deleterious effects. In fact, both the conformational change (p.L339P) and the overall charge change (p.L339R) of the polypeptides cause clinical severe phenotypes simultaneously. Leu339 residue should have some function importance for IDS activity.

4.2.3.3 Mutations at Leu339 residue affect substrate binding

The nine putative active site residues: Asp45, Asn46, Cys84, Arg88, Lys135, His138, Asp334, His335, and Lys347, were proposed to form a loop-like region in an open area of the folded IDS polypeptide (Appendix 10). According to the same 3D structure, Leu339 residue is located in a supercoiled structure adjacent to the active sites region. It seems that the cavity between the active sites region and the supercoiled structure is responsible for substrate binding. Therefore, Leu339 residue is close to a potential substrate binding site. Previous studies suggested that Ser333 to Gly336 may form a region of functional significance located near Leu339 residue [Wilson PJ *et al.*, 1990; Sukegawa K *et al.*, 1995; Isogai K *et al.*, 1998; Froissart R *et al.*, 1998; Li P *et al.*, 1999; Kato T *et al.*, 2005]. This tetrapeptide may be functionally important for substrate binding. Thus, amino acid alterations at Leu339 residue may significantly reduce the substrate binding ability due to conformational change or charge imbalance of this potential substrate binding site.

4.3 Analysis of IDS maturation processing

4.3.1 Active IDS modifications are not completed in lysosomes

The complex processing steps for IDS were studied indirectly. However, not every step was monitored. The location of post-translational processing steps was determined by investigating the activity of the N-terminally truncated IDS in transfected COS-7 cells. The results also showed the significance of the first 33 amino acids at the N-terminal end for IDS localization and functional consequence. This N-terminally truncated IDS was successfully expressed in COS-7 cells and in the cell-free systems. In the absence of signal sequence and the subsequence 24 base pairs, mRNA of this modified IDS was relatively stable as the wild-type (Figure 3.13 C). In the absence of the signal peptide and the subsequence 8 amino acids, the precursors of this modified IDS were quite stable (Figure 3.16). Therefore, the first 99 bp mRNA sequences of *IDS* are not significant for mRNA stability. And the first 33 amino acids are not essential for maintaining protein stability during processing.

In vitro measurement of IDS activity for the N-terminally truncated proteins revealed that no detectable active IDS was found in the transfected cell lysate. All the IDS proteins without signal peptides should be mistargeted and secreted by exocytosis after translation. Interestingly, these proteins without entering the lysosomes were found active in the cell-free medium. Therefore, lacking the 33 amino acids at the N-terminal end does not affect maturation of the polypeptides. The processing steps do not depend on the presence of the first 33 amino acids. The results also indicated that the maturation steps for active IDS were completed before the polypeptides entered the lysosomes. These 33 amino acids containing the signal peptide may only serve for lysosomal targeting. The N-terminal proteolytic

processing should be the final step completed inside the lysosomes.

The results showed that all the processing steps necessary for active IDS are pre-lysosomal. These modifications include the conversion of Cys84 residue in the catalytic domain [Villani GRD *et al.*, 2000; Landgrebe J *et al.*, 2003]. Since the secreted IDS of the N-terminally truncated form was active, the conversion of Cys84 into Cα-formylglycine should be completed. This modification probably started as an early post-translational processing or as a co-translational event in the endoplasmic reticulum [Schmidt B *et al.*, 1995; Dierks T *et al.*, 1997, 1998; Marquordt C *et al.*, 2003; Preusser-Kunze A *et al.*, 2005]. In addition, the results of endocytosis study confirmed that glycosylation was preformed. The secreted IDS was successfully imported into the untransfected COS-7 cells probably by mannose-6-phosphate receptor-mediated cellular uptaking [Millat G *et al.*, 1997b]. It suggested that IDS have already been glycosylated and the oligosaccharide chains were modified in the endoplasmic reticulum [Parkinson-Lawrence E *et al.*, 2005].

4.3.2 C-terminal proteolysis is essential for active IDS

The results also suggested that the C-terminal proteolytsis is partially pre-lysosomal [Schmidt B *et al.*, 1995]. The N-terminally truncated IDS synthesized by the cell-free systems were in precursor form without C-terminal processing. These precursors were inactive which indicated that C-terminal proteolysis determine the catalytic activity of IDS. As mentioned in Chapter 4.2.3.3, there is an open cavity for substrate binding above the loop-like region of active sites residues. Uncleaved C-terminal end may occupy this cavity and bury the active domains inside the folded polypeptides. The results of transient expression showed that IDS

without entering the lysosomes remained only 60 % of wild-type activity. Therefore, incomplete C-terminal cleavages may occur. Probably, the internal cleavage step of the 55 kDa mature form to create the 45-44 kDa second mature form is completed inside the lysosomes. In the absence of the 45-44 kDa mature form, the overall IDS activity may be reduced.

Further investigations of the C-terminal processing will be useful. The second proteolytic cleavage site can be determined for comparing the catalytic activity between the two mature forms of IDS. Since mature IDS have parts of N-terminal and C-terminal sequences removed during processing, fusion proteins of IDS cannot be produced for easy detection. In order to monitor the entire process, specific immunoblot analysis may serve as better visualizing method for transient expression in COS-7 cells. The anti-IDS polyclonal antibodies can specifically stain the IDS proteins for accurate detection in the cell extracts [Parkinson EJ *et al.*, 2004; Parkinson-Lawrence E *et al.*, 2005; Pena O *et al.* 2005]. And the experiments should be performed at different time points within the first 48-hour upon transfection because IDS will be completely processed thereafter [Millat G *et al.*, 1997b].

4.3.3 Functional role of glycosylation during IDS processing

As a lysosomal enzyme, the destination of IDS should be the lysosomes which provide suitable pH for IDS activity. However, previous study proved that it is not the potential N-glycosylation sites responsible for lysosomal targeting of IDS [Millat G *et al.*, 1997a]. Tunicamycin was used to produce unglycosylated IDS precursor to examine its functional consequence [Millat G *et al.*, 1997b]. In this study, the functional role of glycosylation was determined by synthesizing unglycosylated IDS

in the cell-free expression systems. Both the methods could allow us to investigate the cumulative effects of lacking glycosylation on processing and enzyme activity.

It is suggested that the N-oligosaccharide chains formed by glycosylation is responsible for quality control process that ensures correct protein folding [Parkinson-Lawrence E *et al.*, 2005]. Misfolded precursors are degraded in the endoplasmic reticulum [Tikkanen R *et al.*, 1995]. In this study, the cell-free expression systems showed stable protein production of IDS precursors but all the proteins were inactive. These IDS polypeptides had the C-terminal end uncleaved. Therefore, unglycosylated IDS precursors were not able to process C-terminal proteolysis. As mentioned in Chapter 4.3.2, IDS without C-terminal cleavages resulted in enzyme deficiency. The results confirmed that glycosylation of IDS is crucial for IDS folding, processing and catalytic activity.

4.4 Analysis of cell-free expression systems

To date, cell-free *in vitro* expression experiments are rarely applied. It is because the strategies of protein maturations are vary in different enzymes. This is the first study using cell-free systems to synthesize IDS proteins. In the study, different extracts from three different cell types were tried to compare their performance on producing active IDS. However, the results revealed that catalytically active IDS was difficult to be synthesized by any of the cell-free systems. In fact, both the bacterial-based and wheat germ-based systems have good protein expression. It indicated that the sulfatase activity probably depends on normal secondary modifications including formylglycine generation, glycosylation and C-terminal proteolysis. The overall performance of the difference systems is summarized in Table 3.4.

4.4.1 Microbial systems using *E.coli* cell extracts: insoluble IDS precursors

Among the three different types of cell-free expression systems, the bacterial systems were constantly expressing the human proteins. Both *E.coli*-based systems from different manufacturers successfully expressed the IDS precursors. *E.coli* has alternative FGly-generating system to modify cysteine residues for their sulfatases. However, all the precursors were inactive and likely to be trapped in the inclusion bodies. IDS precursors are long polypeptides which have the potential to form partially folded intermediates and aggregated in the bacterial system. The presence of GroE supplements did not assist the systems to form soluble IDS proteins. Indeed, the prokaryotic system is greatly different from eukaryotic system by lacking other

post-translational modifications such as glycosylation [Landgrebe J *et al.*, 2003]. Therefore, the other eukaryotic systems used in this study were expected to improve the situations.

4.4.2 Plant system using wheat germ extracts: soluble IDS precursors

For the wheat germ-based system from Roche, soluble IDS precursors were successfully expressed. The longer incubation time lasted for 24 hours may allow slow folding of the polypeptides. However, the IDS proteins were inactive probably due to the un-modified Cys84 residue and lack of glycosylation. From the SDS-PAGE analysis, the precursor IDS was expressed as one single band with approximately 60 kDa instead of 76 or 90 kDa. The molecular weight revealed that the precursors were unglycosylated without further processing [Millat G *et al.*, 1997b]. This also explained why the soluble IDS failed in entering the COS-7 cells during the endocytosis study. Cellular uptake of IDS is either by cell-cell interaction or M6P receptor-mediated endocytosis [Millat G *et al.*, 1997b].

4.4.3 Mammalian system using rabbit reticulocytes extracts: undetectable

The mammalian system has been successfully used for cell-free expression of active human IDUA and P53 [Keeling KM and Bedwell DM. 2002]. As an open system, microsomal vesicles were added into the rabbit reticulocytes-based expression system to aid post-translational modifications. Only the positive control luciferase was successfully synthesized with activity. The system did not yield any

catalytically active IDS and no detectable IDS protein was visualized even by staining His-tagged proteins. The maturation processing steps of human IDS is much more complex than in human IDUA, human P53 and firefly luciferase. Both the proteolytic cleavages and glycosylation of IDS may be completed with the aid of microsomal vesicles. However, modification of Cys84 residue probably becomes the limiting factor for active IDS synthesis in the mammalian cell-free system.

The FGly generating enzyme (FGE) is a soluble enzyme located in the endoplasmic reticulum for modifying cysteine residues of sulfatases. For microsomal vesicles, all the surface mRNA was stripped. It indicated that the uselessness of adding microsomal vesicles to produce active sulfatases. In addition, protein transfection of the IDS products back into mammalian cells is not practical. It is because FGE only acts on newly synthesized and unfolded sulfatase polypeptides co- or post-translationally [Baenziger JU, 2003; Preusser-Kunze A *et al.*, 2005]. Therefore, future application of the cell-free system should better co-express the IDS or other sulfatases with wild-type FGE. Previous study revealed that co-expression of the *FGE* gene and sulfatases cDNA can increase up-to 50-fold of sulfatase activities in cultured cells [Baenziger JU, 2003].

4.5 Role of transfecting IDS constructs

Bone marrow transplantation may be suitable in mildly affected Hunter disease patients only. For severe patients, recent researches attempt to use enzyme replacement therapy as alternative treatment [Muenzer J *et al.*, 2002]. However, these MPS II patients are also suffered from mental retardation. One of the difficulties for enzyme replacement therapy is that the enzyme cannot enter the nervous system easily. Millat's group suggested that IDS enzymes can be imported into human cells efficiently by cell–to-cell contact [Millat G *et al.*, 1997b]. As some macrophagic cells can enter the central nervous system, long-term transfection of these cells may express normal IDS enzyme and transfer to the nerve cells ultimately. Previous study suggested that the transfection can be performed in the hematopoietic stem cells which will then differentiate into the macrophagic cells [Krivit W *et al.*, 1995]. More recently, muscle electro-gene transfer showed a tenfold increase in IDS activity in a knockout mouse model [Tomanin R *et al.*, 2002]. These approaches may serve as promising gene therapy for MPS II.

4.6 Conclusion

Molecular investigations of a particular mutation could contribute to a better understanding of the disease causing mechanisms. In this study, transient expression confirmed that it is the novel missense mutation p.L339P responsible for clinically severe MPS II. The amino acids alteration cause deficient enzyme activity but its deleterious effect does not affect mRNA stability or protein stability. The hypothetic causality of the mutation may be due to conformational changes. And it is suggested that Leu339 residue may be functionally important for substrate binding. The results also suggested that active IDS is processed pre-lysosomally and glycosylation is essential for IDS processing and catalytic activity. Finally, with the identified mutations, carrier detection and prenatal diagnosis were performed efficiently among the patients' families.

Appendix 1 Multiple alignments of IDS with other human sulfatases

```
IDS_HUMAN    ---MPPPRTGRGLLWLG---LVLSSVCVALGS----ET-QANSTTDALNVLLIIVDDLR-
ARSE_HUMAN   ---MLHLHHSCLCFRSW--LPAMLAVLLSLAP----SASSDISAS-RPNILLLMADDLGI
ARSF_HUMAN   ----------MRPRRP--LVFMSLVCALLNT----WPGHTGCMTTRPNIVLIMVDDLGI
ARSD_HUMAN   MRSAARRGRAAPAARDS--LPVLLFLCLLLKT----CEPKTANAF-KPNILLIMADDLGT
GA6S_HUMAN   ----------MAAVVA--ATRWWQLLLVLSA----AGMGASGAPQPPNILLLLMDDMGW
STS_HUMAN    --------------MP--LRKMKIPFLLLFF----LWEAESHAASRPNIILVMADDLGI
ARSA_HUMAN   ----------------------MGAPRSLLLA----LAAGLAVAR-PPNIVLIFADDLGY
ARSB_HUMAN   ---MGPRGAASLPRGPGPRRLLLPVVLPLLLLLLLAPPGSGAGASRPPHLVFLLADDLGW
SUL2_HUMAN   --MGPPSLVLCLLSATV--FSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQ--
GL6S_HUMAN   -MRLLPLAPGRLRRGSPRHLPSCSPALLLLVLGGCLGVFGVAAGTRRPNVVLLLTDDQ--
SUL1_HUMAN   ---MKYSCCALVLAVLG--TELLGSLCSTVRSPRFRGRIQQERKNIRPNIILVLTDDQ--
                                          *                **  *

IDS_HUMAN    PSLGCYGDKLVRSPNIDQLASHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTRLYDFNS-
ARSE_HUMAN   GDIGCYGNNTMRTPNIDRLAEDGVKLTQHISAASLCTPSRAAFLTGRYPVRSGMVSSIGY
ARSF_HUMAN   GDLGCYGNDTMRTPHIDRLAREGVRLTQHISAASLCSPSRSAFLTGRYPIRSGMVSSGNR
ARSD_HUMAN   GDLGCYGNNTLRTPNIDQLAEEGVRLTQHLAAAPLCTPSRAAFLTGRHSFRSGMDASNGY
GA6S_HUMAN   GDLGVYGEPSRETPNLDRMAAEGLLFPNFYSANPLCSPSRAALLTGRLPIRNGFYTTNAH
STS_HUMAN    GDPGCYGNKTIRTPNIDRLASGGVKLTQHLAASPLCTPSRAAFMTGRYPVRSGMASWSRT
ARSA_HUMAN   GDLGCYGHPSSTTPNLDQLAAGGLRFTDFYVPVSLCTPSRAALLTGRLPVRMGMYPGVLV
ARSB_HUMAN   NDVGFHGSRIR-TPHLDALAAGGVLLDN-YYTQPLCTPSRSQLLTGRYQIRTGLQHQIIW
SUL2_HUMAN   -D-VELGSMQVMNKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKY-VHNHNTYTNN-
GL6S_HUMAN   -D-EVLGGMTPLKKTKALIGEMGMTFSSAYVPSALCCPSRASILTGKYPHNHHVVNNTLE
SUL1_HUMAN   -D-VELGSLQVMNKTRKIMEHGGATFINAFVTTPMCCPSRSSMLTGKY-VHNHNVYTNN-
              *                   *

IDS_HUMAN    --------YWR--VHAGNFSTIPQYFKEN-GYVTMSVGK-----------------VFH
ARSE_HUMAN   RV--LQWTGASGGLPTNE-TTFAKILKEK-GYATGLIGK-----------------WH
ARSF_HUMAN   RV--IQNLAVPAGLPLNE-TTLAALLKKQ-GYSTGLIGK-----------------WH
ARSD_HUMAN   RA--LQWNAGSGGLPENE-TTFARILQQH-GYATGLIGK-----------------WH
GA6S_HUMAN   ARNAYTPQEIVGGIPDSE-QLLPELLKKA-GYVSKIVGK-----------------WH
STS_HUMAN    GV--FLFTASSGGLPTDE-ITFAKLLKDQ-GYSTALIGK-----------------WH
ARSA_HUMAN   P-------SSRGGLPLEE-VTVAEVLAAR-GYLTGMAGK-----------------WH
ARSB_HUMAN   P-------CQPSCVPLDE-KLLPQLLKEA-GYTTHMVGK-----------------WH
SUL2_HUMAN   -----ENCSSPSWQAQHESRTFAVYLNST-GYRTAFFGKYLNEY------NGSYVPPGWK
GL6S_HUMAN   G-----NCSSKSWQKIQEPNTFPAILRSMCGYQTFFAGKYLNEYGAPDAGGLEHVPLGWS
```

```
SUL1_HUMAN  -----ENCSSPSWQAMHEPRTFAVYLNNT-GYRTAFFGKYLNEY------NGSYIPPGWR
                          *       *       *

IDS_HUMAN   PGISSNHTDDSP---YSWS-FPPYHPSSE-------------------------KYEN
ARSE_HUMAN  LGLNCE--SASD---HCHH--PLHHGFEH----------------F------YGM-PFSL
ARSF_HUMAN  QGLNCD--SRSD---QCHH--PYNYGFDY----------------Y------YGM-PFTL
ARSD_HUMAN  QGVNCA--SRGD---HCHH--PLNHGFDY----------------F------YGM-PFTL
GA6S_HUMAN  LGHRPQ--FH----------PLKHGFDE-------------------------------
STS_HUMAN   LGMSCH--SKTD---FCHH--PLHHGFNY----------------F------YGI-SLTN
ARSA_HUMAN  LGVGPE--GAFL---------PPHQGFHR-------------------------------
ARSB_HUMAN  LGMYRK----------ECL--PTRRGFDT-------------------------------
SUL2_HUMAN  EWVGLLKNSRFYNYTLCRNGVKEKHGSDYSKDYLTDLITNDSVSFFRTSKKMYPHRPVLM
GL6S_HUMAN  YWYALEKNSKYYNYTLSINGKARKHGENYSVDYLTDVLAN--VS-LDFLDYKSNFEPFFM
SUL1_HUMAN  EWLGLIKNSRFYNYTVCRNGIKEKHGFDYAKDYFTDLITNESINYFKMSKRMYPHRPVMM
                                    *

IDS_HUMAN   TKTCRGPDGE-----LHANLLCPVDVLDVP---------EGTLP--DKQSTEQAIQLLEK
ARSE_HUMAN  MGDCARWELSEKRVNLEQKLNFLFQVLALVALTLVAGKLTHLIP-VSWMPVIWSALSAVL
ARSF_HUMAN  VDSCWPDPSRNTELAFESQLWLCVQLVAIAILTLTFGKLSGWVS-VPWLLIFSMILFIFL
ARSD_HUMAN  TNDCDPGRPPEVDAALRAQLWGYTQFLALGILTLAAGQTCGFFS-VSARAVTGMAGVGCL
GA6S_HUMAN  ----------------------------------------WFG--SPNCHFGPYDNKAR
STS_HUMAN   LRDCKPGEGSVFTTGFKRLVFLPLQIVGVTLLTLAALNCLGLLH-VPLGVFFSLLFLAAL
ARSA_HUMAN  ----------------------------------------FLG-IPYSHDQGPCQNLTC
ARSB_HUMAN  ----------------------------------------YFG----YLLGSEDYYSHE
SUL2_HUMAN  VISHAAPHGPEDSAPQYSRLFPNASQHITPSYNYAPNPDKHWIMRYTGPMKPIHMEFTNM
GL6S_HUMAN  MIATPAPHSPWTAAPQYQKAFQNVFAPRNKNFNIHGTNKHWLIRQAKTPMTNSSIQFLDN
SUL1_HUMAN  VISHAAPHGPEDSAPQFSKLYPNASQHITPSYNYAPNMDKHWIMQYTGPMLPIHMEFTNI

IDS_HUMAN   MKTSASPFFLAVGYHKPHIPFRYPKEFQKLYPLENITLAPDPEVPDG-LPPVAYNP-WMD
ARSE_HUMAN  LLASSYFVGALIVH-ADCFLMRNHTITEQPMCFQRTTPLILQEVASF-LKRNKHGP--FL
ARSF_HUMAN  LGYAWFSSHTSPLY-WDCLLMRGHEITEQPMKAERAGSIMVKEAISF-LERHSKET--FL
ARSD_HUMAN  FFISWYSSFGFVRR-WNCILMRNHDVTEQPMVLEKTASLMLKEAVSY-IERHKHGP--FL
GA6S_HUMAN  PNIPVYRDWEMVGR--YYEEFPIN----LKTGEANLTQIYLQEALDFIKRQARHHP--FF
STS_HUMAN   ILTLFLGFLHYFRP-LNCFMMRNYEIIQQPMSYDNLTQRLTVEAAQF-IQRNTETP--FL
ARSA_HUMAN  FPPATPCDGGCDQGLVPIPLLANLSVEAQPPWLPGLEARYMAFAHDLMADAQRQDR-PFF
ARSB_HUMAN  RCT--LIDALNVTRC--ALDFRDGEEVATGYKNMYSTNIFTKRAIALITNHPPEKP--LF
SUL2_HUMAN  LQRKRLQTLMSVDDSMETIYNMLVETGELDNTYIVYTADHGYHIGQFGLVKGKSMPYEFD
```

```
GL6S_HUMAN   AFRKRWQTLLSVDDLVEKLVKRLEFTGELNNTYIFYTSDNGYHTGQFSLPIDKRQLYEFD
SUL1_HUMAN   LQRKRLQTLMSVDDSVERLYNMLVETGELENTYIIYTADHGYHIGQFGLVKGKSMPYDFD

IDS_HUMAN    IRQREDVQALNIS----VPYGPIPVDFQRKIRQSYF--ASVSYLDTQ-VGRLLS---ALD
ARSE_HUMAN   LFVSFLH----VH----IPLITMENFLGKSLHGLYG--DNVEEMDWM-VGRILD---TLD
ARSF_HUMAN   LFFSFLH----VH----TPLPTTDDFTGTSKHGLYG--DNVEEMDSM-VGKILD---AID
ARSD_HUMAN   LFLSLLH----VH----IPLVTTSAFLGKSQHGLYG--DNVEEMDWL-IGKVLN---AIE
GA6S_HUMAN   LYWAVDA----TH----APVYASKPFLGTSQRGRYG--DAVREIDDS-IGKILE---LLQ
STS_HUMAN    LVLSYLH----VH----TALFSSKDFAGKSQHGVYG--DAVEEMDWS-VGQILN---LLD
ARSA_HUMAN   LYYASHH----TH----YPQFSGQSFAERSGRGPFG--DSLMELDAA-VGTLMT---AIG
ARSB_HUMAN   LYLALQS----VHEPLQVPEEYLKPYDFIQDKNRHHYAGMVSLMDEA-VGNVTA---ALK
SUL2_HUMAN   IRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDIAGL-DIPADMDGKSILKLLDTERPVN
GL6S_HUMAN   IKVPLLVRGPGIKPNQTSKMLVANIDLGPTILDIAGYDLNKTQMDGMSLLPILR------
SUL1_HUMAN   IRVPFFIRGPSVEPGSIVPQIVLNIDLAPTILDIAGL-DTPPDVDGKSVLKLLDPEKPGN

IDS_HUMAN    DLQLANSTIIA---FTSDHG------------------WALG----------------E
ARSE_HUMAN   VEGLSNSTLIY---FTSDHG----------------GSLENQLG-------------NTQ
ARSF_HUMAN   DFGLRNNTLVY---FTSDHG----------------GHLEARRG-------------HAQ
ARSD_HUMAN   DNGLKNSTFTY---FTSDHG----------------GHLEARDG-------------HSQ
GA6S_HUMAN   DLHVADNTFVF---FTSDNG----------------AALISAP---------------EQ
STS_HUMAN    ELRLANDTLIY---FTSDQG----------------AHVEEVSSK------------GEI
ARSA_HUMAN   DLGLLEETLVI---FTADNG------------------PETMRMS--------------R
ARSB_HUMAN   SSGLWNNTVFI---FSTDNG------------------GQTLAG----------------
SUL2_HUMAN   RFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEENFLPKYQRVKDLCQRAEYQTACEQ
GL6S_HUMAN   G----ASNLTWRSDVLVEYQ-------------------------------------------
SUL1_HUMAN   RFRTNKKAKIWRDTFLVERGKFLRKKEESSKNIQQSNHLPKYERVKELCQQARYQTACEQ
                          *    *

IDS_HUMAN    HG-------------EWAKYSN--FDVATH------VPLIFY-------------VPGRTA
ARSE_HUMAN   YG-GW------NGIYKGGKGMG-GWEGGIR------VPGIFR-------------WPGVLP
ARSF_HUMAN   LG-GW------NGIYKGGKGMG-GWEGGIR------VPGIVR-------------WPGKVP
ARSD_HUMAN   LG-GW------NGIYKGGKGMG-GWEGGIR------VPGIFH-------------WPGVLP
GA6S_HUMAN   GG--------SNGPFLCGKQTT--FEGGMR------EPALAW-------------WPGHVT
STS_HUMAN    HG-GS------NGIYKGGK-AN-NWEGGIR------VPGILR-------------WPRVIQ
ARSA_HUMAN   GG--------CSGLLRCGKGTT--YEGGVR------EPALAF-------------WPGHIA
ARSB_HUMAN   -G--------NNWPLRGRKWSL--WEGGVRG-----VGFVAS-------------PLLKQ
```

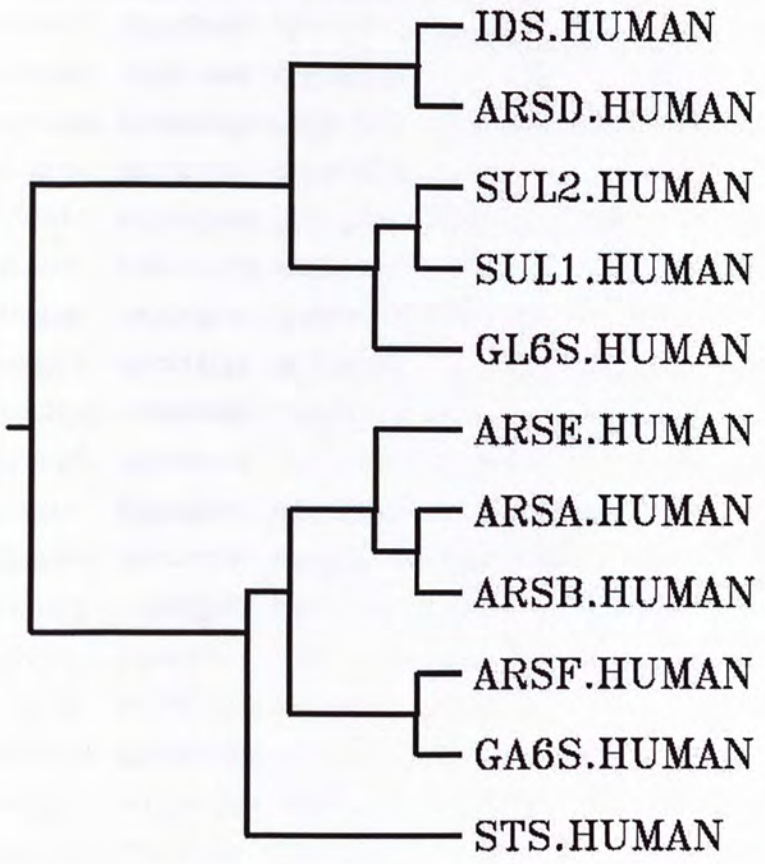
```
SUL2_HUMAN  LGQKWQCVEDATGKLKLHKCKGPMRLGGSR-ALSNLVPKYYGQGSEACTCDSGDYKLSLA
GL6S_HUMAN  ------------------------------------------------------------
SUL1_HUMAN  PGQKWQCIEDTSGKLRIHKCKGPSDLLTVRQSTRNLYARGFHDKDKECSCRESGYRASRS
             *                    *

IDS_HUMAN   SLP----EAG------EKLFPYLD--PFDSASQLME---------------PGRQ--SMD
ARSE_HUMAN  AGR----VIG-------EPTSLMD--VFPTVVRLAG-------G-------EVPQ---DR
ARSF_HUMAN  AGR----LIK-------EPTSLMD--ILPTVASVSG-------G-------SLPQ---DR
ARSD_HUMAN  AGR----VIG-------EPTSLMD--VFPTVVQLVG-------G-------EVPQ---DR
GA6S_HUMAN  AGQ--VS---------HQLGSIMD--LFTTSLALAG---------------LTPP--SDR
STS_HUMAN   AGQ----KID-------EPTSNMD--IFPTVAKLAG-------A-------PLPE---DR
ARSA_HUMAN  PGVT-----------HELASSLD--LLPTLAALAG--------------APLPN----V
ARSB_HUMAN  KGVK----NR-------ELIHISD--WLPTLVKLAR---------------GHTN--GTK
SUL2_HUMAN  GRR-KKLFKK-----KYKASYVRSRSIRSVAIEVDGRVYHVGLG-----DAAQPRNLTKR
GL6S_HUMAN  ----------------------------------G-------------------------
SUL1_HUMAN  QRKSQRQFLRNQGTPKYKPRFVHTRQTRSLSVEFEGEIYDINLEEEEELQVLQPRNIAKR

IDS_HUMAN   LVEL----VSLFPTLAG---------LAGLQVP-------PRCPVPSFHVELCREG---K
ARSE_HUMAN  VIDG----QDLLPLLLG--------TAQHSDHE----FLMHYCERFLHAARWHQRD---R
ARSF_HUMAN  VIDG----RDLMPLLQG--------NVRHSEHE----FLFHYCGSYLHAVRWIPKDD--S
ARSD_HUMAN  VIDG----HSLVPLLQG--------AEARSAHE----FLFHYCGQHLHAARWHQKD---S
GA6S_HUMAN  AIDG----LNLLPTLLQ---------GRLMDRP-----IFYYRGDTLMAATLGQHK--AH
STS_HUMAN   IIDG----RDLMPLLEG--------KSQRSDHE----FLFHYCNAYLNAVRWHPQN---S
ARSA_HUMAN  TLDG----FDLSPLLLG---------TGKSPRQ-----SLFFYPSYPDEVRGVFAV---R
ARSB_HUMAN  PLDG----FDVWKTISE---------GSPSPRI----ELLHNIDPNFVDSSPCPRN---S
SUL2_HUMAN  HWPG--APEDQDDKDGGDFSGTGGLPDYSAANP---IKVTHRCYILENDTVQCDLDLYKS
GL6S_HUMAN  --EG----RNVTDPTCP--------SLSPGVSQ---CFPDCVCEDAYNNTYACVRT---M
SUL1_HUMAN  HDEGHKGPRDLQASSGG-NRGRMLADSSNAVGPPTTVRVTHKCFILPNDSIHCERELYQS

IDS_HUMAN   NLLK--HFR--------FR-----DLEEDPYLPGNPRELIAYS-QYPRPSDIPQWNSDKP
ARSE_HUMAN  GTMWKVHFV-----TPVFQ------PEGAGACYGRKVCPCFGE-KVVHHDP-PLLFDLSR
ARSF_HUMAN  GSVWKAHYV-----TPVFQ------PPASGGCYVTSLCRCFGE-QVTYHNP-PLLFDLSR
ARSD_HUMAN  GSVWKVHYT-----TPQFH------PEERGLLTAEASAHAEWG-GVTHHRP-PLLFDLSR
GA6S_HUMAN  FWTW--TNS-----WENFR-------QGIDFCPGQNVSGVTTH-NLEDHTKLPLIFHLGR
STS_HUMAN   TSIWKAFFF-----TPNFN------PVGSNGCFATHVCFCFGS-YVTHHDP-PLLFDISK
ARSA_HUMAN  TGKYKAHFF----------------TQGSAHSDTTADPACHASSSLTAHEP-PLLYDLSK
```

```
ARSB_HUMAN  MAPAKDDSS---------------LPEYSAFNTSVHAAIRHGNWKLLTGYP-GCGYWFPP
SUL2_HUMAN  LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCH-KISYHTQ-HKGRLKHR
GL6S_HUMAN  SALWNLQYC------EFDD------QEVFVEVYNLTADPDQIT-NIAKTID-PELLGKMN
SUL1_HUMAN  ARAWKDHKAYIDKEIEALQDKIKNLREVRGHLKRRKPEECSCS-KQSYYNK-EKGVKKQE

IDS_HUMAN   S-LKDIKIMGYSIRT---IDYRYTVWVGFNPDEFLANFSDIH------AGELYFVDSDPL
ARSE_HUMAN  D-PSETHILTPASEP------VFYQVMERVQQAVWEHQRTLS------PVPLQLDRLGNI
ARSF_HUMAN  D-PSESTPLTPATEP------LYDFVIKKVANALKEHQETIV------PVTYQLSELN-Q
ARSD_HUMAN  D-PSEARPLTPDSEP------LYHAVIARVGAAVSEHRQTLS------PVPQQFSMSNIL
GA6S_HUMAN  D-PGERFPLSFASAE-------YQEALSRITSVVQQHQEALV------PAQPQLNVCNWA
STS_HUMAN   D-PRERNPLTPASEP------RFYEILKVMQEAADRHTQTLP------EVPDQFSWNNFL
ARSA_HUMAN  D-PGENYNLLGGVAG------ATPEVLQALKQLQLLKAQLDA------AVTFGPSQVARG
ARSB_HUMAN  PSQYNVSEIPSSDPP---TKTLWLFDIDRDPEERHDLSREYP-----HIVTKLLSRLQFY
SUL2_HUMAN  --GSSLHPFRKGLQEKD-KVWLLREQ-KRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQH
GL6S_HUMAN  ---YRLMMLQSCSGPTCRTPGVFDPGYRFDPRLMFSNRGSVR-------TR-RFSKHLL-
SUL1_HUMAN  KLKSHLHPFKEAAQEVDSKLQLFKENNRRRKKERKEKRRQRKGEECSLPGLTCFTHDNNH

IDS_HUMAN   QDHN--MYND----SQGGDLFQLLMP----------------------------------
ARSE_HUMAN  WRP---WLQP----CCGPFPL-CWCLREDDPQ--------------------------
ARSF_HUMAN  GRT---WLKP----CCGVFPF-CLCDKEEEVSQPRGPNEKR-------------------
ARSD_HUMAN  WKP---WLQP----CCGHFPF-CSCHEDGDGTP--------------------------
GA6S_HUMAN  VMN---WAPP----GCEKLGK-CLTPPESIPKKCLWSH---------------------
STS_HUMAN   WKP---WLQL----CCPSTGLSCQCDREKQDKRLSR-------------------------
ARSA_HUMAN  EDP---ALQICCHPGCTPRPACCHCPDPHA-----------------------------
ARSB_HUMAN  HKHSVPVYFPAQDPRCDPKATGVWGPWM-------------------------------
SUL2_HUMAN  WQTAPFWTLGPFCACTSANNNTYWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNA
GL6S_HUMAN  ------------------------------------------------------------
SUL1_HUMAN  WQTAPFWNLGSFCACTSSNNNTYWCLRTVNETHNFLFCEFATGFLEYFDMNTDPYQLTNT

IDS_HUMAN   ------------------------------------------------------------
ARSE_HUMAN  ------------------------------------------------------------
ARSF_HUMAN  ------------------------------------------------------------
ARSD_HUMAN  ------------------------------------------------------------
GA6S_HUMAN  ------------------------------------------------------------
STS_HUMAN   ------------------------------------------------------------
```

```
ARSA_HUMAN  ------------------------------------------------------------
ARSB_HUMAN  ------------------------------------------------------------
SUL2_HUMAN  VNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYEQYRQFQRRKWPEMK
GL6S_HUMAN  ------------------------------------------------------------
SUL1_HUMAN  VHTVERGILNQLHVQLMELRSCQGYKQCNPRPKNLDVGNKDGGSYDLHR----------

IDS_HUMAN   ----------------
ARSE_HUMAN  ----------------
ARSF_HUMAN  ----------------
ARSD_HUMAN  ----------------
GA6S_HUMAN  ----------------
STS_HUMAN   ----------------
ARSA_HUMAN  ----------------
ARSB_HUMAN  ----------------
SUL2_HUMAN  RPSSKSLGQLWEGWEG
GL6S_HUMAN  ----------------
SUL1_HUMAN  -------GQLWDGWEG
```

Appendix 1 The multiple alignments of human sulfatases' amino acid sequences were done by Clustal W. IDS has strong sequence homology with other human sulfatases. Highly conserved residues are shared in yellow. The conserved pentapeptide for sulfatase activity is boxed. Amino acids identical in at least ten sulfatases are starred on the bottom line. The Leu339 residue of IDS is shared in pink. IDS: iduronate-2-sulfatase; ARSE: arylsulfatase E; ARSF: arylsulfatase F; ARSD: arylsulfatase D; GA6S: N-acetylgalactosamine-6-sulfatase; STS: steroid sulfatase or arylsulfatase C; ARSA: galactose 3-sufatase or arylsulfatase A; ARSB: N-acetylgalactosamine 4-sulfatase or arylsulfatase B; SUL2: extracellular sulfatase Sulf-2; GL6S: N-acetylglucosamine-6-sulfatase and; SUL1: extracellular sulfatase Sulf-1.

Appendix 2 Gene tree of human sulfatases



Appendix 2 It is suggested that the human sulfatases comprise an evolutionarily related family of genes that arose by gene duplication and divergent evolution. IDS: iduronate-2-sulfatase; ARSD: arylsulfatase D; SUL2: extracellular sulfatase Sulf-2; SUL1: extracellular sulfatase Sulf-1; GL6S: N-acetylglucosamine-6-sulfatase; ARSE: arylsulfatase E; ARSA: galactose 3-sufatase or arylsulfatase A; ARSB: N-acetylgalactosamine 4-sulfatase or arylsulfatase B; ARSF: arylsulfatase F; GA6S: N-acetylgalactosamine-6-sulfatase and; STS: steroid sulfatase or arylsulfatase C.

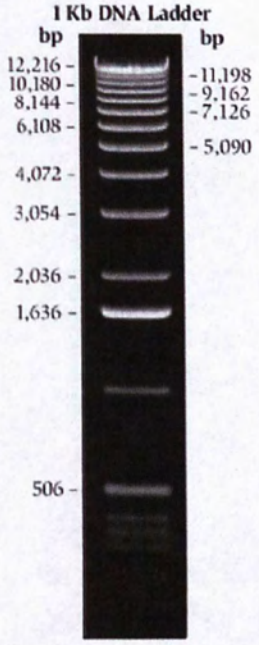
Appendix 3 cDNA sequence of *IDS* gene

```
   1 atgccgcca ccccggaccg gccgaggcct tctctggctg ggtctggttc tgagctccgt
  61 ctgcgtcgcc ctcggatccg aaacgcaggc caactcgacc acagatgctc tgaacgttct
 121 tctcatcatc gtggatgacc tgcgcccctc cctgggctgt tatggggata agctggtgag
 181 gtccccaaat attgaccaac tggcatccca cagcctcctc ttccagaatg cctttgcgca
 241 gcaagcagtg tgcgccccga gccgcgtttc tttcctcact ggcaggagac ctgacaccac
 301 ccgcctgtac gacttcaact cctactggag ggtgcacgct ggaaacttct ccaccatccc
 361 ccagtacttc aaggagaatg gctatgtgac catgtcggtg ggaaaagtct ttcaccctgg
 421 gatatcttct aaccataccg atgattctcc gtatagctgg tcttttccac cttatcatcc
 481 ttcctctgag aagtatgaaa acactaagac atgtcgaggg ccagatggag aactccatgc
 541 caacctgctt tgccctgtgg atgtgctgga tgttcccgag ggcaccttgc ctgacaaaca
 601 gagcactgag caagccatac agttgttgga aaagatgaaa acgtcagcca gtccttttctt
 661 cctggccgtt gggtatcata agccacacat ccccttcaga taccccaagg aatttcagaa
 721 gttgtatccc ttggagaaca tcaccctggc ccccgatccc gaggtccctg atggcctacc
 781 ccctgtggcc tacaacccct ggatggacat caggcaacgg gaagacgtcc aagccttaaa
 841 catcagtgtg ccgtatggtc caattcctgt ggactttcag cggaaaatcc gccagagcta
 901 ctttgcctct gtgtcatatt tggatacaca ggtcggccgc ctcttgagtg ctttggacga
 961 tcttcagctg gccaacagca ccatcattgc atttacctcg gatcatgggt gggct[cta]gg
1021 tgaacatgga gaatgggcca aatacagcaa ttttgatgtt gctacccatg ttcccctgat
1081 attctatgtt cctggaagga cggcttcact tccggaggca ggcgagaagc ttttccctta
1141 cctcgaccct tttgattccg cctcacagtt gatggagcca ggcaggcaat ccatggacct
1201 tgtggaactt gtgtctcttt ttcccacgct ggctggactt gcaggactgc aggttccacc
1261 tcgctgcccc gttccttcat ttcacgttga gctgtgcaga gaaggcaaga accttctgaa
1321 gcattttcga ttccgtgact tggaagagga tccgtacctc cctggtaatc cccgtgaact
1381 gattgcctat agccagtatc cccggccttc agacatccct cagtggaatt ctgacaagcc
1441 gagtttaaaa gatataaaga tcatgggcta ttccatacgc accatagact ataggtatac
1501 tgtgtgggtt ggcttcaatc ctgatgaatt tctagctaac ttttctgaca tccatgcagg
1561 ggaactgtat tttgtggatt ctgacccatt gcaggatcac aatatgtata atgattccca
1621 aggtggagat cttttccagt tgttgatgcc ttga
```

Appendix 3 The cDNA sequence of wild-type *IDS* gene was adapted from NM_000202 (NCBI). The N-terminally truncated *IDS* constructs containing all the coding sequence except the first 99 base pairs (underlined). Codon 339, i.e. CTA, was bolded and boxed.

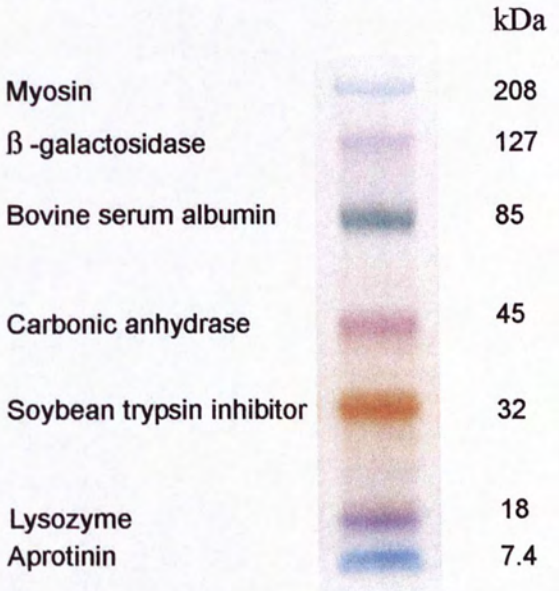
Appendix 4 Details of markers

Appendix 4.1 1kb DNA Ladder

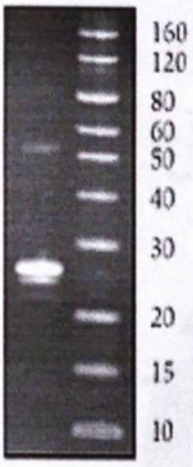


0.9 µg/lane

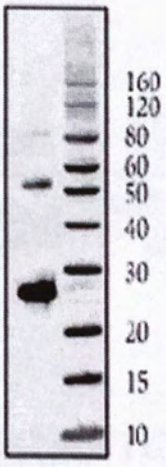
Appendix 4.2 Kaleidoscope Prestained Standards



Appendix 4.3 BenchMark His-tagged Standards



A. Stained with InVision™ His-tag In-gel Stain



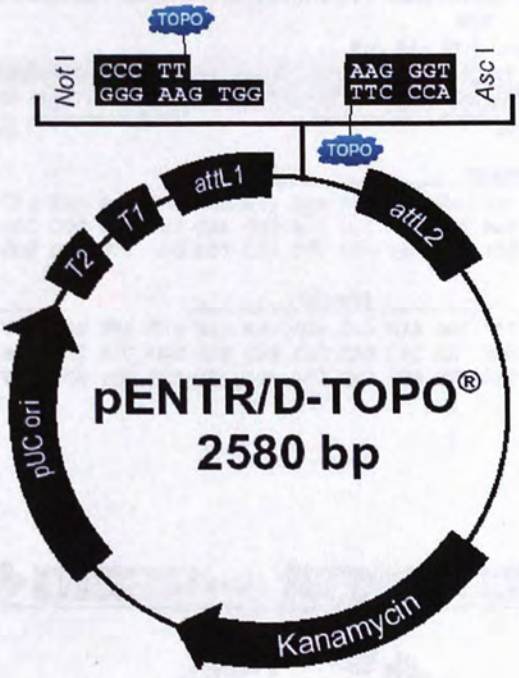
Lane 1: 5 µl of cell lysate with expressed His-tagged sequence

Lane 2: 5 µl of the BenchMark™ His-tagged Standard

B. Blotted onto nitrocellulose, followed by detection with the Anti-His(C-term) Antibody and visualized using WesternBreeze™ Chemiluminescent Kit

Appendix 5 Vector information of pENTR/D-TOPO

Appendix 5.1 Map

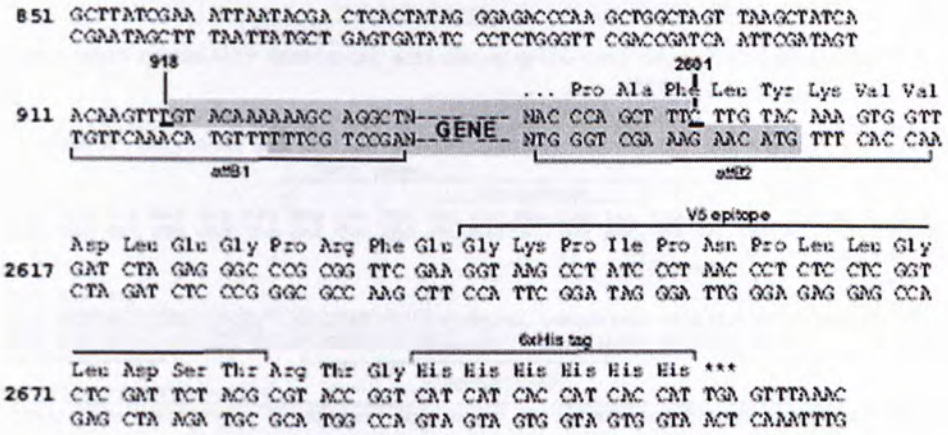


Elements: rnB T2 transcription termination sequence: bases 268-295
rrnB T1 transcription termination sequence: bases 427-470
M13 forward (-20) priming site: bases 537-552
attL1: bases 569-668 (complementary strand)
TOPO recognition site 1: bases 680-684
Overhang: bases 685-688
TOPO recognition site 2: bases 689-693
attL2: bases 705-804
M13 reverse priming site: bases 845-861
Kanamycin resistance gene: bases 974-1783
pUC origin: bases 1904-2577

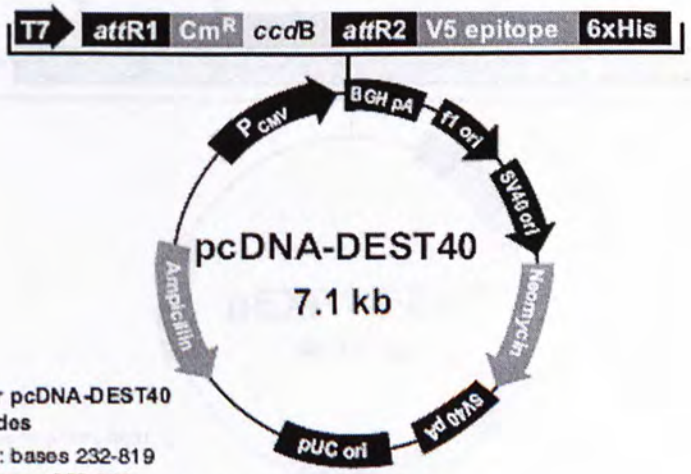
Appendix 5 pENTR/D-TOPO vector is linearized with topoisomerase I covalently bound to each 3'-phosphate at both ends. Then, a GTGG overhanging on the 5'-end and a blunt end on the 3'-end are generated. The four-nucleotide overhang anneals to the 5'CACC sequence of the PCR products so that the insert is in a desired 5'→3' orientation. The vector contains the attL recombination regions for efficient swapping of the insert with other Gateway destination vectors; kanamycin resistance gene for selection in E.coli.

Appendix 6 Vector information of pcDNA-DEST40

Appendix 6.1 Multiple cloning site



Appendix 6.2 Map



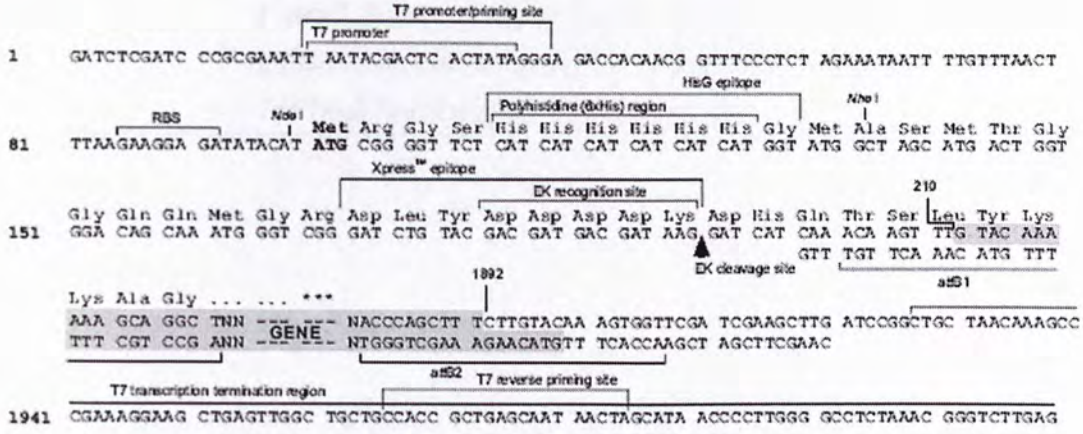
Comments for pcDNA-DEST40
7143 nucleotides

CMV promoter: bases 232-819
T7 promoter: bases 863-882
attR1 recombination site: bases 911-1035
Chloramphenicol resistance gene: bases 1144-1803
ccdB gene: bases 2145-2450
attR2 recombination site: bases 2491-2615
V5 epitope: bases 2641-2682
6xHis tag: bases 2692-2709
BGH polyadenylation region: bases 2735-2962
f1 origin: bases 3008-3436
SV40 early promoter and origin: bases 3463-3771
Neomycin resistance ORF: bases 3846-4640
SV40 early polyadenylation region: bases 4816-4946
pUC origin: bases 5329-6002
Ampicillin (*bla*) resistance ORF: bases 6147-7007 (complementary strand)
bla promoter: bases 7008-7106 (complementary strand)

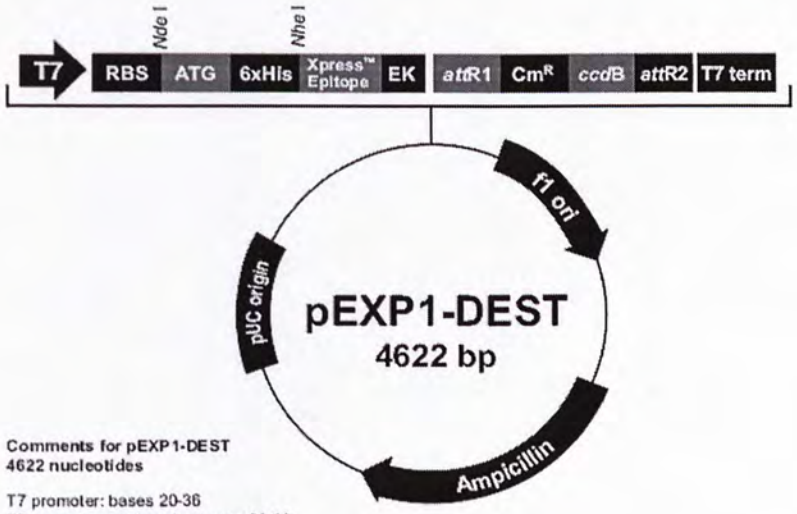
Appendix 6 The human CMV promoter: for high-level expression in mammalian cells; bovine growth hormone (BGH) polyadenylation sequence: for proper termination of the transcript; the attR recombination regions for efficient swapping of insert with pENTR/D-TOPO.

Appendix 7 Vector information of pEXP1-DEST

Appendix 7.1 Recombination region



Appendix 7.2 Map



Comments for pEXP1-DEST
4622 nucleotides

T7 promoter: bases 20-36
T7 promoter priming site: bases 20-39
Ribosome binding site: bases 85-92
Initiation ATG: bases 100-102
Polyhistidine (6xHis) region: bases 112-129
Xpress™ epitope: bases 169-192
Enterokinase (EK) recognition site: bases 178-192
*attR*1 site: bases 202-326
Chloramphenicol resistance gene (Cm^R): bases 435-1094
ccdB gene: bases 1436-1741
*attR*2 site: bases 1782-1906
T7 reverse priming site: bases 1966-1985
T7 transcription termination region: bases 1927-2056
f1 origin: bases 2127-2582
bla promoter: bases 2669-2767
Ampicillin resistance gene: bases 2768-3628
pUC origin: 3773-4446

Appendix 7 The vector contains transcription / translation regulatory elements, including bacteriophage T7 RNA polymerase promoter, prokaryotic Shine-Dalgarno ribosome binding site (RBS) and T7 terminator. 6×His sequence enables rapid detection and purification of the fusion protein. It contains the attR recombination regions for efficient swapping of insert with pENTR/D-TOPO.

Appendix 8 Genotype of bacterial cells

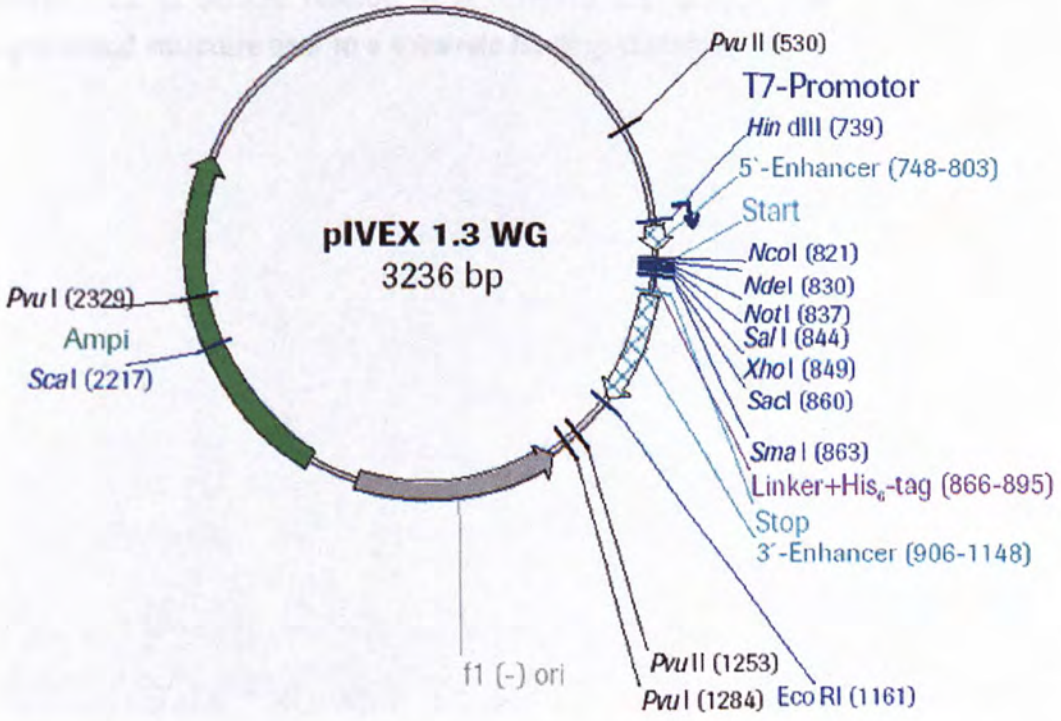
Competent cells	Genotype
TOP10	F⁻*mcr*A△(*mrr-hsd*RMS-*mcr*BC)Ø80*lac*Z△M15 △*lac*X74*deo*R*rec*A1*ara*D139△(*ara-leu*) *7697gal*U*gal*K*rps*L(Str^{R})*end*A1*nup*G

Appendix 9 Vector information of pIVEX_1.3_WG

Appendix 9.1 Multiple cloning site

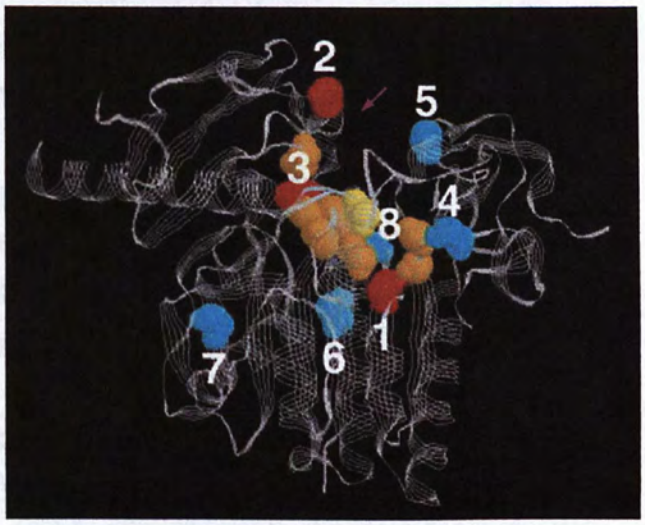
```
                          T7-Promotor              HindIII
701 TTACGCCAAG CTCATTAATA CGACTCACTA TAGGCCTAAG CTTACAAATA
    AATGCGGTTC GAGTAATTAT GCTGAGTGAT ATCCGGATTC GAATGTTTAT

                        5`-Enhancer
751 CTCCCCCACA ACAGCTTACA ATACTCCCCC ACACAGCTTA CAAATACTCC
    GAGGGGGTGT TGTCGAATGT TATGAGGGGG TGTGTCGAAT GTTTATGAGG

                       Start
                      NcoI    NdeI      NotI    SalI
801 CCCACAACAG CTTGTCGAAC CATGGCACAT ATGAGCGGCC GCGTCGACTC
    GGGTGTTGTC GAACAGCTTG GTACCGTGTA TACTCGCCGG CGCAGCTGAG

  XhoI SacI  SmaI          Linker+His6-tag          Stop
851 GAGCGAGCTC CCGGGGGGGG TTCTCATCAT CATCATCATC ATTAATAAGG
    CTCGCTCGAG GGCCCCCCCC AAGAGTAGTA GTAGTAGTAG TAATTATTCC

                       3'-Enhancer
901 TACCCAGCTC TTCTGGTTTG GTTTGGACCT CTGGTCCTGC AACTTGAGGT
    ATGGGTCGAG AAGACCAAAC CAAACCTGGA GACCAGGACG TTGAACTCCA
```

Appendix 9.2 Map



Appendix 9 The vector contains transcription / translation regulatory elements, including bacteriophage T7 RNA polymerase promoter, 5′ and 3′ enhancers. 6×His sequence enables rapid detection and purification of the fusion protein.

Appendix 10 Predicted 3D structure of IDS polypeptide



Appendix 10 The predicted 3D structure of IDS polypeptides was adapted from Kato T *et al.*, 2005. The orange spheres represent the putative active sites. The numbered spheres were used to indicate the altered residues in Kato's study. Red sphere no.2 is Ser336 residue. It is believed that Leu339 residue is located in a supercoiled structure near to a substrate binding site(arrowed).

Electronic-database and computing system

URLs for data presented are as follows:

http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=309900 [for MPS II in OMIM Database]

http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=5360215 [for *IDS* cDNA sequence in NCBI]

http://www.hgmd.cf.ac.uk/hgmd0.html [for *IDS* in Human Gene Mutation Database]

http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi [for primer design using Primer3]

http://dove.embl-heidelberg.de/Blast2e/ [for multiple aligment of human sulfatases using CLUSTALW]

http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx [for hairpin prediction using OligoAnalyzer 3.0]

http://tools.neb.com/NEBcutter2/index.php [for restriction site prediction using NEBcutter V2.0]

Bibliography

Annella T, Daniele A and Di Natale P. Heterogeneity of DNA and RNA in Hunter patients. *Human Genetics* Oct 92 (4): 350-352. 1993.

Baenziger JU. A major step on the road to understanding a unique posttranslational modification and its role in a genetic disease. *Cell* 113: 421-422. 2003.

Betton JM. Rapid translation system (RTS): A promise alternative for recombinant protein production. *Current Protein and Peptide Science* 4 (1): 73-78. 2003

Birnboim HC and Doly J. A rapid alkaline lysis procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 7: 1513-1522. 1979.

Birot AM, Bouton O, Froissart R, Maire I and Bozon D. IDS gene-pseudogene exchange responsible for an intragenic deletion in a hunter patient. *Human Mutation* 8 (1): 44-50. 1996.

Bondeson ML, Dahl N, Malmgren H, Kleijer WJ, Tonnesen T, Carlberg BM and

Pettersson U. Inversion of the IDS gene resulting from recombination with IDS-related sequences is a common cause of the Hunter syndrome. Human Molecular Genetics 4: 615-621. 1995.

Bonuccelli G, Di Natale P, Corsolini F, Villani GRD, Regis S and Filocamo M. The effect of four mutations on the expression of iduronate-2-sulfatase in mucopolysaccharidosis type II. Biochimica et Biophysica Acta 1537: 233-238. 2001.

Braun SE, Aronovitch EL, Anderson RA, Crotty PL, Mclvor RS and Whitley CB. Metabolic correction and cross-correction of mucopolysaccharidosis type II (Hunter's syndrome) by retroviral-mediated gene transfer and expression of human iduronate-2-sulfatase. Proceedings of the National *Academy of Sciences of the United States* 90: 11830-11834. 1993.

Bunge S, Steglich C, Beck M et al. Mutation analysis of the iduronate-2-sulfatase gene in patients with mucopolysaccharidosis type II (Hunter syndrome). Human Molecular Genetics 1: 335-339. 1992.

Chang JH, Lin SP, Lin SC, Tseng KL, Li CL, Chuang CK and Lee-Chen GJ. Expression studies of mutation underlying Taiwanese Hunter syndrome (mucopolysaccharidosis type II). Human Genetics 116: 160-166. 2005.

Ciccarone V, Chuy Schifferi, K Pichet JP, Hawley-Nelson P, Evans K, Roy L, and Bennett S. Lipofectamine 2000 reagent for rapid, efficient transfection of eukaryotic cells. Focus 21: 54-55. 1999.

Coppa GV, Gabrielli O, Zampini L, Jetzequel AM, Miniero R, Busca A, et al. Bone marrow transplantation in Hunter syndrome. Journal of Inherited Metabolic Disease 18: 91-92. 1995.

Coonrod A, Li FQ and Horwitz M. On the mechanism of DNA transfection: efficient gene transfer without viruses. Gene Therapy 4 (12): 1313-1321. 1997.

Cooper DN and Krawczak M. Human gene mutation. Oxford: BIOS Scientific Publishers. 1993.

Cudry S, Tigaud I, Froissart R, Bonnet V, Maire I and Bozon D. MPS II in females: molecular basis of two different cases. Journal of Medical Genetics 37(10): e29.

2000.

Cuer M, Barnier A, de La Salmoniere P, Durand G and Seta N. Fluorimetric measurement of plasma α–L-fucosidase activity with a centrifugal analyzer: reference values in a healthy French adult population. Clinical Chemistry 46: 560-576. 2000.

Daniele A, Faust CJ, Herman GE, Di Natale P and Ballobio A. Cloning and characterization of the cDNA for the murine iduronate sulfatase gene. Genomics 16 (3): 755-757. 1993.

Dierks T, Schmidt B and von Figura K. Conversion of cysteine to formylglycine: a protein modification in the endoplasmic reticulum. Proceedings of the National Academy of Sciences of USA 94: 11963-11968. 1997.

Dierks T, Lecca M, Schmidt B and von Figura K. conversion of cysteine to formylglycine in eukaryotic sulfatases occurs by a common mechanism in the endoplasmic reticulum. FEBS Letters. 423: 61-65. 1998.

Dierks T, Lecca MR, Schlotterhose P, Schimdt B and von Figura K. Sequence determinants directing conversion of cysteine to formylglycine in eukaryotic sulfatases. EMBO Journal 18: 2084-2091. 1999.

Felgner PL, Gadek TR, Holm M, Roman R, Chan HW, Wenz M, Northrop JP, Ringold GM, and Danielsen M. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proceedings of the National Academy of Sciences of USA 84 (21): 7413-7417. 1987.

Filocamo M, Bonuccelli G, Corsolini F, Mazzotti R, Cusano R and Gatti R. Molecular analysis of 40 Italian patients with mucopolysaccharidosis type II: New mutations in the iduronate-2-sulfatases (IDS) gene. Human Mutation 18 (2): 164-165. 2001

Flomen RH, Green PM, Bentley DR and Giannelli F. Determination of the organization of coding sequences within the iduronate sulfatase (IDS) gene. Human Molecular Genetics 2: 5-10. 1993.

Froissart R, Millat G, Mathieu M, Bozon D and Maire I. Processing of iduronate-2-sulphatase in human fibroblasts. Biochemical Journal 309 (Pt 2): 425-430. 1995.

Froissart R, Maire I, Millat G, Cudry s, Birot AM, Bonnet V, Bouton O, Bozon D. Identification of iduronate sulfatase gene alterations in 70 unrelated Hunter patients. Clinical Genetics May 53 (5): 362-368. 1998

Gamper HB, Parekh H, Rice MC, Bruner M, Youkey H and Kmiec EB. The DNA strand of chimeric RNA/DNA oligonucleotides can direct gene repair/conversion activity in mammalian and plant cell-free extracts. Nucleic Acids Research 28 (21): 4332-4339. 2000.

Gartler SM and Riggs AD. Mammalian X-chromosome inactivation. Annual Review of Genetics 17: 155-190. 1983.

Gershon H, Ghirlando R, Cuttman SB and Minsky A. Mode of formation and structural features of DNA-cationic liposome complexes used for transfection. Biochemistry 32 (28): 7143-7151. 1993.

Goldenfum SL, Young E, Michelakakis H, Tsagarakis S and Winchester B. Mutation analysis in 20 patients with Hunter disease. Human Mutation 7: 76-78. 1996.

Graumann K and Premstaller A. Manufacturing of recombinant therapeutic proteins in microbial systems. Biotechnology Journal 1 (2): 164-186. 2006

Hong YT, Yu SS, Kim JM, Lee K, Na YS, Whitley CB, Sugimoto Y and Kim SY. Construction of a high efficiency retroviral vector for gene therapy of Hunter's syndrome. Journal of Gene Medicine 5: 18-29. 2003.

Hopwood JJ, Bunge S, Morris CP, Wilson PJ, Steglich C, Beck M, Schwinger E and Gai A. Molecular basis of mucopolysaccharidosis type II: Mutations in the iduronate-2-sulphatase gene. Human Mutation 2 (6): 435-442. 1993.

Isogai K, Sukegawa K, Tomatsu S, Fukao T. Song XQ, Yamada Y, Fukuda S, Orii T and Kondo N. Mutation analysis in the iduronate-2-sulphatase gene in 43 Japanese patients with mucopolysaccharidosis type II (Hunter disease). Journal of Inherited Metabolic Disease 21: 60-70. 1998.

Jonsson JJ, Aronovich EL, Braun SE and Whitley CB. Molecular diagnosis of mucopolysaccharidosis type II (Hunter syndrome) by automated sequencing and computer-assisted interpretations towards mutation mapping of the iduronate-2-sulfatase gene. Amerian Journal of Human Genetics Mar 56 (3):

597-607. 1995.

Kato T, Kato Z, Kuratsubo I, Tanaka N, Ishigami T, Kajihara JI, Sukegawa-Hayasaka K, Orii K, Isogai K, Fukao T, Shimozawa N, Orii T, Kondo N and Suzuki Y. Mutational and structural analysis of Japanese patients with mucopolysaccharidosis type. Journal of Human Genetics Aug 50 (8): 395-402. 2005.

Keeling KM and Bedwell DM. Clinically relevant aminoglycosides can suppress disease-associated premature stop mutations in the IDUA and P53 cDNAs in a mammalian translation system. Journal of Molecular Medicine 80: 367-376. 2002.

Keulemans JLM, Sinigerska I, Garritsen VH, Huijmans JGM, Voznyi YV, van Diggelen OP and Kleijer WJ. Prenatal diagnosis of the Hunter syndrome and the introduction of a new fluorimetric enzyme assay. Prenatal Diagnosis 22: 1016-1021. 2002.

Kim CH, Hwang HZ, Song SM, Paik KH, Kwon EY, Moon KB, Yoon JH, Han CK and Jin DK. Mutational spectrum of the iduronate-2-sulfatase gene in 25 unrelated Korean Hunter syndrome patients: identification of 13 novel mutations. Human Mutation 21 (4): 449-450. 2003.

Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Research 15: 8125-8148. 1987.

Krivit W, Sung JH, Shapiro EG and Lockman LA. Microglia: the effector cell for reconstitution of the central nervous system following bone marrow transplantation for lysosomal and peroxisomal storage diseases. Cell Transplant 4: 385-392. 1995.

Landgrebe J, Dierks T, Schmidt B and von Figura K. The human SUMF 1 gene, required for posttranslational sulfatase modification, defines a new gene family which is conserved from pro- to eukaryotes. Gene 316: 47-56. 2003.

Landy A. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. Annual Review of Biochemistry 58: 913-949. 1989.

Lesley SA, Brow MA and Burgess RR. Use of in vitro protein synthesis from polymerase chain reaction-generated templates to study interaction of Escherichia coli transcription factors with core RNA polymerase and for epitope mapping of monoclonal antibodies. Journal of Biological Chemistry 266 (4): 2632-2638. 1991.

Li P, Bellows AB and Thompson JN. Molecular basis of iduronate-2-sulphatase gene mutations in patients with mucopolysaccharidosis type II (Hunter syndrome). Journal of Medical Genetics 36: 21-27. 1999.

Liu S, Li L, Fu J, Zhong C and Lu G. The detection of the frequent mutations of iduronate-2-sulphatase gene in mucopolysaccharidosis type patients in Chinese. Zhonghua Yi Xue Yi Chuan Xue Za Zhi Jun 19 (3): 243-245. 2002.

Lyon MF. Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190: 372-373. 1961.

Malmgren H, Carlberg BM, Pettersson U and Bondeson ML. Identification of an alternative transcript from the human iduronate-2-sulfatase (IDS) gene. Genomics Sep 1; 29 (1): 291-293. 1995.

Marquordt C, Fang Q, Will E, Peng J, von Figura K and Dierks T. Posttranslational modification of serine to formylglycine in bacterial sulfatase. Recognition of the modification motif by the iron-sulfur protein. AtsB. Journal of Biological Chemistry 278: 2212-2218. 2003.

McClelland M and Nelson M. Effect of site-specific methylation on DNA modification methyltransferases and restriction endoncleases. Nucleic Acids Research 20 Supplement: 2145-2157. 1992.

McGuire JBJ, James TJ, Imber CJ, St. Peter SD, Friend PJ and Taylor RP. Optimisation of an enzymatic method for β-galactosidase. Clinica Chimica Acta 326: 123-129. 2002.

Millat G, Froissart R, Maire I and Bozon D. Characterization of iduronate sulphatase mutants affecting N-glycosylation sites and the cysteine-84 residue. Biochemical Journal 326: 243-247. 1997a.

Millat G, Froissart R, Maire I and Bozon D. IDS transfer from overexpressing cells to IDS-deficient cells. Experimental Cell Research 230: 362-367. 1997b.

Millat G, Froissart F, Cudry S, Bonnet V, Maire I and Bozon D. COS cell expression studies of P86L, P86R, P480L and P480Q Hunter's disease-causing mutations. Biochimica et Biophysica Acta Mar 5; 1406 (2): 214-218. 1998.

Moreira da Silva I, Froissart R, Marques dos Santos H, Caseiro C, Maire I and Bozon D. Molecular basis of mucopolysaccharidosis type II in Portugal: identification of four novel mutations. Clinical Genetics 60: 316-318. 2001.

Muenzer J, Lamsa JC, Garcia A, Dacosta J, Garcia J and Treco DA. Enzyme replacement therapy in mucopolysaccharidosis type II (Hunter syndrome): a preliminary report. Acta Paediatrica Supplementum 91 (439): 98-99. 2002.

Nakatsu T, Ichiyama S, Hiratake J, Saldanha A, Kobashi N, Sakata K and Kato H. Structural basis for the spectral difference in luciferase bioluminescence. Nature Mar 440 (7082): 372-376. 2006.

Nelson J. Incidence of the mucopolysaccharidosis in Northern Ireland. Human Genetics 101 (3): 355-358. 1997.

Neufeld EF and Muenzer J. The mucopolysaccharidoses; In: Scriver CR, Beaudet AL, Sly WS, Valle D, Childs B, Kinzler KW and Vogelstein B (eds). The metabolic and molecular bases of inherited disease (8th edn). McGraw-Hill New York V.3 p 3421-3452. 2001.

Park JG and Chapman VM. CpG island promoter region methylation patterns of the inactive-X-chromosome hypoxanthine phosphoribosyltransferase (Hprt) gene. Molecular Cell Biology 14: 7975-7983. 1994.

Parkinson EJ, Muller V, Hopwood JJ and Brooks DA. Iduronate-2-sulphatase protein detection in plasma from mucopolysaccharidosis type II patients. Molecular Genetics and Metabolism Jan; 81 (1): 58-64. 2004.

Parkinson-Lawrence E, Turner C, Hopwood JJ and Brooks D. Analysis of normal and mutant iduronate-2-sulphatase conformation. Biochemical Journal 386: 395-400. 2005.

Patharkar OR and Cushman JC. A stress-induced calcium-dependent protein kinase fromMesembryanthemum Crystallinum phosphorylates a two-component pseudo-response regulator. The Plant Journal 24 (5): 679-691. 2000.

Pena O, Sosa A, Echeverri O, Saenz H and Barrera LA. Production of polyclonal antibodies to protein iduronate-2- sulphate sulphatase (IDS) and development of a detection system for human recombinant IDS. Biomedica Jun; 25 (2): 181-188.

Spanish. 2005.

Preusser-Kunze A, Mariappan M, Schmidt B, Gande SL, Mutenda K, Wenzel D, von Figura K and Dierks T. Molecular characterization of the human Calpha-formylglycine-generating enzyme. Journal of Biological Chemistry 289 (15): 14900-14910. 2005.

Ranthmann M, Bunge S, Beck M, Kresse H, Tylki-Szymanska A, and Gal A. Mucopolysaccharidosis type II (Hunter Syndrome): Mutation "hot spots" in the iduronate-2-sulfatase gene. American Journal of Human Genetics 59: 1202-1209. 1996.

Ricci V, Filocamo M, Regis S, Corsolini F, Stroppiano M, Di Duca M and Gatti R. Expression studies of two novel in CIS-mutations identified in an intermediate case of Hunter syndrome. American Journal of Medical Genetics 120A: 84-87. 2003.

Sanger F, Nicklen S and Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of USA 74 (12): 5463-5467. 1977.

Schmidt B, Selmer T, Ingendoh A and von Figura K. A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. Cell Jul 28; 82 (2): 271-278. 1995.

Shuman S. Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. Journal of Biological Chemistry 23, 269 (51): 32678-32684. 1994.

Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeysinghe S, Krawczak M and Cooper DN. Human gene mutation database (HGMD): 2003 update. Human Mutation 21: 577-581. 2003.

Stroncek DF, Hubel A, Shankar RA, Burger SR, Pan D, McCullough J and Whitley CB. Retroviral transduction and expansion of peripheral blood lymphocytes for the treatment of mucopolysaccharidosis type II, Hunter's syndrome. Transfusion 39 (4): 343-50. 1999.

Sukegawa K, Matsuzaki T, Fukuda S, Masuno M, Fukao T, Kokuryu M, et al. Brother/sister siblings affected with Hunter disease: evidence for skewed X chromosomes inactivation. Clincial Genetics 53: 96-101. 1998.

Sukegawa K, Nakamura H, Kato Z, Tomatsu S, Montano AM, Fukao T, Toietta G, Tortora P, Orii T and Kondo N. Biochemical and structural analysis of missense mutations in N-acetylgalactosamine-6-sulfate sulfatase causing mucopolysaccharidosis IVA phenotypes. Human Molecular Genetics 9: 1283-1290. 2000.

Sukegawa K, Tomatsu S, Fukao T, Iwata H, Song XQ, Yamada Y, Fukuda S, Isogai K and Orii T. Mucopolysaccharidosis type II (Hunter disease): identification and characterization of eight point mutations in the iduronate-2-sulfatase gene in Japanese patients. Human Mutation 6 (2): 136-143. 1995.

Thomas GH. "Pseudodeficiencies" of lysosomal hydrolases. American Journal of Human Genetics 54 (6): 934-940. 1994.

Tikkanen R, Enomaa N, Riikonen A, Ikonen E and Pelthonen L. Intracellular sorting of aspartylglucosaminidase: the role of N-linked oligosaccharides and evidence of Man-6-P-independent lysosomal targeting. DNA and cell biology 14 (4): 305-312. 1995.

Timms KM, Lu F, Shen Y, Pierson CA, Muzny DM, Gu Y and Nelson DL. 130 Kb of DNA sequence reveals two new genes and a regional duplication distal to the human iduronate-2-sulfate sulfatase locus. Genome Research 5: 71-78. 1995.

Tomanin R, Friso A, Alba S, Piller Puicher E, Mennuni C, La Monica N, Hortelano G, Zacchello F and Scarpa M. Non-viral transfer approaches for the gene therapy of mucopolysaccharidosis type II (Hunter syndrome). Acta Paediatrica Nov supplement 439 (91): 100-104. 2002.

Tomatsu S, Orii KO, Bi Y, Gutierrez MA, Nishioka T, Yamaguchi S, Kondo N, Orii T, Noguchi A and Sly WS. General implications for CpG hot spot mutations: methylation patterns of the human iduronate-2-sulfatase gene locus. Human Mutation 23: 590-598. 2004.

Villani GRD, Daniele A, Balzano N and Natale PD. Expression of five iduronate-2-sulfatase site-directed mutations. Biochimica et Biophysica Acta 1501: 71-80. 2000.

Voloshin AM and Swartz JR. Efficient and scalable method for scaling up cell free protein synthesis in batch mode. Biotechnology and Bioengineering 91 (4): 516-521.

2005.

von Heijne G.. A new method for predicting signal sequence cleavage sites. Nucleic Acids Research 14: 4683-4690. 1986.

Voznyi YV, Keulemans JLM and van Diggelen OP. A fluorimetric enzyme assay for the diagnosis of MPS II (Hunter disease). Journal of Inherited Metabolic Disease 24: 675-680. 2001.

Warkentin PI, Dixon MS Jr, Schafer I, Strandjord SE and Coccia PF. Bone marrow transplantation in Hunter syndrome: a preliminary report. Birth Defects Original Article Series 22 (1): 31-39. 1986.

Whitley CB, Mclvor RS, Aronovich EL, et al. Retroviral-mediated transfer of the iduronate-2-sulfatase gene into lymphocytes for treatment of mild Hunter syndrome (mucopolysaccharidosis type II). Human Gene Therapy 7: 537-549. 1996.

Whittington R and Goa KL. Alglucerase: A review of its therapeutic use in Gaucher's disease. Drugs 44 (7): 72-93. 1992

Wilson PJ, Morris CP, Anson DS, Occhiodoro T, Bielicki J, Clements PR and Hopwood JJ. Hunter syndrome: Isolation of an iduronate-2-sulfatase cDNA clone and analysis of patient DNA. Proceedings of the National Academy of Sciences of USA 87: 8531-8535. 1990.

Wilson PJ, Meaney CA, Hopwood JJ and Morris CP. Sequence of the human iduronate-2-sulfatase (IDS) gene. Genomics 17: 773-775. 1993.

Winchester B, Young E, Geddes S, Genet S, Hurst J et al. Female twin with Hunter disease due to nonrandom inactivation of the X-chromosome: a consequence of twinning. American Journal of Medical Genetics 44: 834-838. 1992

Young ID, Harper PS, Newcombe RG and Archer IM. A clinical and genetic study of Hunter's syndrome. 2. Differences between the mild and severe forms. Journal of Medical Genetics 19: 408-411. 1982.

Zubay G. In vitro synthesis of protein in microbial systems. Annual Review of Genetics 7: 267-287.